This is a digital copy of a book that was preserved for generations on library shelves before it was carefully scanned by Google as part of a project to make the world's books discoverable online.

It has survived long enough for the copyright to expire and the book to enter the public domain. A public domain book is one that was never subject to copyright or whose legal copyright term has expired. Whether a book is in the public domain may vary country to country. Public domain books are our gateways to the past, representing a wealth of history, culture and knowledge that's often difficult to discover.

Marks, notations and other marginalia present in the original volume will appear in this file - a reminder of this book's long journey from the publisher to a library and finally to you.

Usage guidelines

Google is proud to partner with libraries to digitize public domain materials and make them widely accessible. Public domain books belong to the public and we are merely their custodians. Nevertheless, this work is expensive, so in order to keep providing this resource, we have taken steps to prevent abuse by commercial parties, including placing technical restrictions on automated querying.

We also ask that you:

+ *Make non-commercial use of the files* We designed Google Book Search for use by individuals, and we request that you use these files for personal, non-commercial purposes.

+ *Refrain from automated querying* Do not send automated queries of any sort to Google's system: If you are conducting research on machine translation, optical character recognition or other areas where access to a large amount of text is helpful, please contact us. We encourage the use of public domain materials for these purposes and may be able to help.

+ *Maintain attribution* The Google "watermark" you see on each file is essential for informing people about this project and helping them find additional materials through Google Book Search. Please do not remove it.

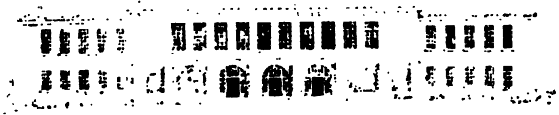
+ *Keep it legal* Whatever your use, remember that you are responsible for ensuring that what you are doing is legal. Do not assume that just because we believe a book is in the public domain for users in the United States, that the work is also in the public domain for users in other countries. Whether a book is still in copyright varies from country to country, and we can't offer guidance on whether any specific use of any specific book is allowed. Please do not assume that a book's appearance in Google Book Search means it can be used in any manner anywhere in the world. Copyright infringement liability can be quite severe.

About Google Book Search

Google's mission is to organize the world's information and to make it universally accessible and useful. Google Book Search helps readers discover the world's books while helping authors and publishers reach new audiences. You can search through the full text of this book on the web at `http://books.google.com/`

Stanford University Libraries
3 6105 006 580 497

LELAND STANFORD JVNIOR VNIVERSITY



SCHOOL OF EDUCATION
LIBRARY

Journal of Educational Research

Official Organ of the Educational Research Association

VOLUME VI

June-December, 1922

PUBLIC SCHOOL PUBLISHING COMPANY
Bloomington, Illinois

Copyright 1922 by
Public School Publishing Company

300525

INDEX

Volume VI, June-December, 1922

Abilities, A Study in Correlation of Allied English, Grace S. Jamison, 241
Ability, Mental, How much does a Teacher Need? W. B. Bliss, 33
Ability to Learn Foreign Languages, Prognosis Tests of, Thos. H. Briggs, 386
Achievement, The Efficiency Quotient as a Measure of, T. L. Torgerson, -25
Achievements of High School and University Students in Certain Tasks in Chemistry, A Comparison of the, S. R. Powers, 332
Acquiring Skill in Teaching, James R. Grant, (R), (Dwight H. Stevenson) -448
Adjustments in California, Will C. Wood, 326
Administrative Use of Intelligence Tests, Some Pitfalls in the, M. R. Trabue, 1
Adults, The Mental Age of, E. A. Lincoln, 133
Adults and Related Problems, Intelligence Quotients of, Walter F. Dearborn, 307
Adults, The Mental Age of, (E), F. N. Freeman, 441
Age Determined by Annual Retests, Changes in Mental Traits with, (R), F. D. Brooks, (V. A. C. Henmon), 72
Algebra, Comparative Validity of the Hotz Scales and the Rugg-Clark Tests in, Eleanore Harris and Frederick S. Breed, 393
Allied English Abilities, A Study in Correlation of, Grace S. Jamison, 241
Alternate Responses, Criticism of Tests Requiring, H. H. Hahn, 236
American Citizenship Course in United States History, (R), Fannie Fern Andrews, (R. M. Tryon), 67
Americanization, Study in; Nationality and School Progress, (R), Riverda Harding Jordan, (Josephine D. Mason), 66
Andrews, Fannie Fern: An American Citizenship Course in United States History, (R), (R. M. Tryon), 67
Annual Re-tests, Changes in Mental Traits with Age Determined by, (R), F. D. Brooks, (V. A. C. Henmon), 72
Another Educational Campaign, J. W. Richardson, 97
Arithmetic, The Psychology of, (R), Edward L. Thorndike, (Cyrus D. Mead), 163
Ashbaugh, E. J.: Some Essentials in School Supervision, 116
Attempts at Test Validation, Raymond Franzen, 145
Attendance and Child Labor, Compulsory School, (R), F. Chester Ensign (C. C. McCracken), 73
Attitudes in Social Education, The Functions of Ideals and, (R), Paul Frederick Voelker (W. J. Osburn), 168

Bagley, W. C.: Professor Terman's Determinism: A Rejoinder, 371
Baldwin, Bird T. and Lorle I. Stecher: Mental Growth Curve

of Normal and Superior Children, (R), (G. M. Ruch), 261
Bamberger, Florence Eilau: The Effect of the Physical Make-up of a book upon Children's Selection, (R), (Harold E. Burtt), 170
Berry, Chas. S.: Classification by Test of Intelligence of Ten Thousand First-Grade Pupils, 185
Betts, G. H.: The New Program of Religious Education, (R), (L. A. Weigle), 265
Binet-Simon and Otis Tests, Comparison of the, S. C. Garrison and James S. Tippett, 42
Binet-Simon Tests, Herring Revision of the, John P. Herring, (R), (P. R. Stevenson), 259
Binet Scale, The Classification of Kindergarten Children for First-Grade by means of the, Chas. D. Dawson, 413
Blackburn, Mary: Montessori Experiments in a large Infants' School, (R), (Luella A. Palmer), 267
Bliss, W. B.: How much Mental Ability does a Teacher Need? -33
Bobbitt, Franklin: Curriculum Making in Los Angeles, (R), (W. W. Charters), 347
Bode, Boyd H.: Fundamentals of Education, (R), (Stephen S. Colvin) -71
Book, William F.: The Intelligence of High School Seniors, (R), (Dorothy H. Yates), 449
Boraas, Julius: Teaching to Think, (R), (N. S. Maddox), 265
Breed, Frederick S. and Eleanore Harris: Comparative Validity of the Hotz Scales and the Rugg-Clark Tests in Algebra, 393
Briggs, Thomas H.: Prognosis Tests of Ability to Learn Foreign Languages, 386
Bright and Dull Pupils, Can Teachers Select, G. F. Verner, 126
Brooks, Fowler Dell: Changes in Mental Traits with Age Determined by annual Re-tests, (R), (V. A. C. Henmon), 72
Buckingham, B. R.: The Principal as Executive, (E), 254
Buckingham, B. R.: Progress and Research, (E), 257
Buckingham, B. R.: A Challenge, (E), 345
Buckingham, B. R.: The Difficulty of Textbooks, (E), 444
Butterworth, Julian E.: Rural School Survey of New York State, (R), (P. R. Stevenson), 450

Calculation and Interpretation of Percentile Ranks, L. L. Thurstone, 225
California, Adjustments in, Will C. Wood, 326
Campaign, Another Educational, J. W. Richardson, 97
Can Teachers Select Bright and Dull Pupils? G. F. Verner, 126
Capps, A. G.: Functions of Tests and Scales as Found in Recent Educational Literature, 204
Challenge, A., B. R. Buckingham, (E), 345
Changes in Mental Traits with age determined by Annual Re-tests, F. D. Brooks, (R), (V. A. C. Henmon), 72
Charters, W. W.: How to Sell at Retail, (R), (B. R. Buckingham), 451
Chemistry, Comparison of the Achievements of High-School

and University Students in Certain Tasks in, S. R. Powers, 332
Child Labor, Compulsory School Attendance and, F. C. Ensign, (R), (C. C. McCracken), 73
Childhood and Children, (R), N. Niemeyer, (A. Sophie Rogers), 169
Children, Classification of Kindergarten for First Grade by means of the Binet Scale, Chas. D. Dawson, 413
Children and Childhood, (R), N. Niemeyer, (A. Sophie Rogers), 169
Children's Selection, The Effect of the Physical Make-up of a book upon, Florence Eilau Bamberger, (R), (Harold E. Burtt), 170
Chronological Age in Decimal Parts of a Year, Determining, H. A. Toops, 438
Citizenship Course in United States History, An American, (R), Fannie Fern Andrews, (R. M. Tryon), 67
City School Superintendent, The Legal Status of the, (R), J. Cayce Morrison, (F. C. Landsittell), 352
City School Finances, The Control of, (R), Geo. W. Frasier, (B. R. Buckingham), 259
Civic Education, David Snedden, (R), (Chas. C. Peters), 350
Clark-Rugg Tests in Algebra, Comparative Validity of the Hotz Scales and the, Eleanore Harris and Frederick S. Breed, 393
Classification of Kindergarten Children for First Grade by Means of the Binet Scale, Chas. D. Dawson, 412
Classification by Test of Intelligence of Ten Thousand First-Grade Pupils, Chas. S. Berry, 185
Co-Efficient, A Data Sheet for the Pearson, Correlation, L. L. Thurstone, 49
Comparative Validity of the Hotz Scales and the Rugg-Clark Tests in Algebra, Eleanore Harris and Frederick S. Breed, 393
Comparison of the Achievements of High School and University Students in Certain Tasks in Chemistry, S. R. Powers, 332
Comparison of the Binet-Simon and Otis Tests, S. C. Garrison and James S. Tippett, 42
Composition, Curriculum Study in Letter-Writing and English, Roy Ivan Johnson, 424
Compulsory School Attendance and Child Labor, F. C. Ensign, (R), (C. C. McCracken), 73
Connersville Course of Study in Elementary Mathematics, (R), G. M. Wilson, (Madge Perrill), 262
Content of the Elementary School Course in Reading, Scientific Determination of the, (R), W. L. Uhl, (C. R. Stone), 449
Control of City School Finances, The, (R), Geo. W. Frasier, (B. R. Buckingham), 259
Correlation of Allied English Abilities, A Study in, Grace S. Jamison, 241
Correlation Co-Efficient, A. Data Sheet for the Pearson, L. L. Thurstone, 49
Course in United States History, An American Citizenship, (R) Fannie Fern Andrews, (R. M Tryon), 67
Course in Reading, Scientific Determination of the Content of the Elementary School, (R), W. L. Uhl, (C. R. Stone), 449
Course of Study in Elementary Mathematics, Connersville, (R),

G. M. Wilson, (Madge Perrill), 262
Courtis, S. A.: Reaction, (E), 159
Criticism of Tests Requiring Alternate Responses, H. H. Hahn, 236
Curricula in the North Central States from 1860 to 1918, The Development of High School, (R), John Elbert Stout, (H. A. Hollister), 162
Curriculum-Making in Los Angeles, Franklin Bobbitt, (R), (W. W. Charters), 347
Curriculum Study in English Composition, Letter-Writing, . Roy Ivan Johnson, 423

Dalton, Laboratory Plan (R), Evelyn Dewey, (Harriet O'Shea), 167
Data Sheet for the Pearson Correlation Co-Efficient, L. L. Thurstone, 49
Dawson, Chas. D.: The Classification of Kindergarten Children for First-Grade by means of the Binet Scale, 412
Dearborn, W. F. and Alexander Inglis: Psychological and Educational Tests in the Public Schools of Winchester, Virginia, (R), (W. W. Coxe), 260
Dearborn, W. F.: Intelligence Quotients of Adults and Related Problems, 307
Democracy and the I.Q., The Psychological Determinist, (E), Lewis M. Terman, 57
Determination of the Content of the Elementary School Course in Reading, Scientific, (R), W. L. Uhl, (C. R. Stone), 449
Determinism: A Rejoinder, Professor Terman's, W. C. Bagley, 371
Determinist, The Psychological: or Democracy and the I.Q., (E), Lewis M. Terman
Development of High-School Curricula in the North Central States from 1860 to 1918, (R), John Elbert Stout, (H. A. Hollister), 162
Developments in Silent Reading Tests, Recent, C. R. Stone, 102
Dewey, Evelyn: The Dalton Laboratory Plan, (R), (Harriet O'Shea), 167
Differential Education as applied to Handicapped Pupils in the Elementary Grades, Theory of, J. E. W. Wallin, 209
Difficulty of Textbooks, (E), B. R. Buckingham, 444
Disability, A Study of Reading and Spelling with Special Reference to, Arthur R. Gates, 12
Drawing, A Measuring Scale for Freehand, (R), L. W. Kline and Gertrude L. Carey, (Wm. H. Vogel), 349
Dull Pupils, Can Teachers Select Bright and, G. F. Verner, 126
Dunn, Fannie: Interest Factors in Primary Reading Material, (R), (Frances Jenkins), 69

Education, Fundamentals of, (R), Boyd H. Bode, (Stephen S. Colvin), 71
Education, Trade Tests in, (R), Herbert A. Toops, (J. Cayce Morrison), 74
Education, The Function of Ideals and Attitudes in Social, (R), Paul Frederick Voelker, (W. J. Osburn), 168
Education as Applied to Handicapped Pupils in the Elementary Grades, Theory of Differential, J. E. W. Wallin, 209
Education, The New Program of Religious, (R), G. H. Betts, (L. A. Weigle), 265
Education, Civic, (R), David Snedden, (Chas. C. Peters), 350

Educational Campaign, Another, J. W. Richardson, 97

Educational Literature, Functions of Tests and Scales as Found in Recent, A. G. Capps, 204

Educational Tests in the Public Schools of Winchester, Virginia, Psychological and (R), W. F. Dearborn and Alexander Inglis, (W. W. Coxe), 260

Effective Study, Training for, (R), F. W. Thomas, (B. H. Bode), 348

Effective Expression, (R), Chas. E. Rhodes, (W. L. Graves), 351

Efficiency Quotient as a measure of Achievement, T. L. Torgerson, 25

Elementary Grades, Theory of Differential Education as Applied to Handicapped Pupils in the, J. E. W. Wallin, 209

Elementary Mathematics, Connersville Course of Study in, (R), G. M. Wilson, (Madge Perrill), 262

Elementary School Course in Reading, Scientific Determination of the Content of the, (R), W. L. Uhl, (C. R. Stone), 449

Empirical Studies in School Reading, (R), James F. Hosic, (S. S. Brooks), 68

English Abilities, A Study in Correlation of Allied, Grace S. Jamison, 241

English Composition, Letter-Writing: A Curriculum Study in, Roy Ivan Johnson, 424

Ensign, Forest Chester: Compulsory School Attendance and Child Labor, (R), (C. C. McCracken), 73

Essentials in School Supervision, Some, E. J. Ashbaugh, 116

Essentials in Place Geography, Minimum, P. R. Stevenson, 300

Executive, The Principal as (E), B. R. Buckingham, 254

Experiments in a Large Infants School, Montessori, (R), Mary Blackburn, (Luella A. Palmer), 267

Expression, Effective, (R), Chas. E. Rhodes, (W. L. Graves), 351

Factors in Primary Reading Material, Interest, (R), Fannie W. Dunn, (Frances Jenkins), 69

Finances, The Control of City School, (R), Geo. W. Frasier, (B. R. Buckingham), 259

Foreign Languages, Prognosis Tests of Ability to Learn, Thos. H. Briggs, 386

Foundations, of Personality, The, Abraham Myerson (R), (Ernest N. Henderson), 70

Franzen, R. H. and F. B. Knight, Textbook Selection (R), (Herman Remmers), 447

Franzen, R. H.: Attempts at Test Validation, 145

Frasier, Geo. W.: The Control of City School Finances, (R), (B. R. Buckingham), 259

Freehand Drawing, A Measuring Scale for, (R), L. W. Kline and Gertrude L. Carey, (William H. Vogel), 349

Freeman, F. N.: The Mental Age of Adults (E), 441

Function of Ideals and Attitudes in Social Education, (R), Paul Frederick Voelker, (W. J. Osburn), 168

Functions of Tests and Scales as Found in Recent Educational Literature, A. G. Capps, 204

Fundamentals of Education, (R), Boyd H. Bode, (Stephen S. Colvin), 71

Garrison, S. C. and James S. Tippett: Comparison of the Binet-Simon and Otis Tests, 42

Gates, Arthur R.: A Study of Reading and Spelling with Special Reference to Disability, 12

Geography, Teaching by Problems, (R), Ehrlich E. Smith, (Douglas C. Ridgley), 75

Geography, Minimum Essentials in Place, P. R. Stevenson, 300

Grade-First, Classification by Test of Intelligence of Ten Thousand, Chas. S. Berry, 185

Grades-Elementary, Theory of Differential Education as Applied to Handicapped Pupils in the, J. E. W. Wallin, 209

Grant, James R.: Acquiring Skill in Teaching, (R), (Dwight H. Stevenson), 448

Growth Curve of Normal and Superior Children, Mental, (R), Bird T. Baldwin, (G. M. Ruch), 261

Handbook of Mental Tests, (R), F. Kuhlmann, (Laura M. Chassell), 165

Handicapped Pupils in the Elementary Grades, Theory of Differential Education as Applied to, J. E. W. Wallin, 209

Hahn, H. H.: Criticism of Tests Requiring Alternate Responses, 236

Harris, Eleanore and Frederick S. Breed: Comparative Validity of the Hotz Scales and the Rugg-Clark Tests in Algebra, 393

Herring, John P.: Herring Revision of the Binet-Simon Tests, (R), (P. R. Stevenson), 259

Herring Revision of the Binet-Simon Tests, (R), John P. Herring, (P. R. Stevenson), 259

High-School Curricula in the North Central States from 1860 to 1918, The Development of, (R), John Elbert Stout, (H. A. Hollister), 162

High-School and University Students in Certain Tasks in Chemistry, A Comparison of the Achievements of, S. R. Powers, 332

High-School Seniors, The Intelligence of, (R), Wm. F. Book, (Dorothy H. Yates), 449

History, An American Citizenship Course in United States, (R), Fannie Fern Andrews, (R. M. Tryon), 67

Hosic, James F.: Empirical Studies in School Reading, (R), (S. S. Brooks), 68

Hotz Scales and the Rugg-Clark Tests in Algebra, Comparative Validity of the, Eleanore Harris and Frederick S. Breed, 393

Ideals and Attitudes in Social Education, The Function of, (R), Paul Frederick Voelker, (W. J. Osburn), 168

Infants School, Montessori Experiments in a large, (R), Mary Blackburn, (Luella A. Palmer), 267

Inglis, Alexander, W. F. Dearborn and others: Psychological and Educational Tests in the Public Schools of Winchester, Virginia, (R), (W. W. Coxe), 260

Intelligence Tests, Some Pitfalls in the Administrative use of, M. R. Trabue, 1

Intelligence of Ten Thousand First-Grade Pupils, Classification by Test, Chas. S. Berry, 185

Intelligence Quotients of Adults and Related Problems, Walter F. Dearborn, 307

Intelligence of High School Seniors, The, (R), Wm. F. Book, (Dorothy H. Yates), 449

Interest Factors in Primary Reading Material, (R), Fannie W. Dunn, (Frances Jenkins), 69
Interpretation of Percentile Ranks, Calculation and, L. L. Thurstone, 225

Jamison, Grace S.: A Study in Correlation of Allied English Abilities, 241
Johnson, Roy Ivan: Letter-Writing: A Curriculum Study in English Composition, 423
Jordan, Riverda Harding: Nationality and School Progress: A Study in Americanization, (R), (Josephine D. Mason), 66

Kindergarten Children for First-Grade by means of the Binet-Scale, The Classification of, Chas. D. Dawson, 412
Kline, L. Ward and Gertrude L. Carey: A Measuring Scale for Freehand Drawing, (R), (Wm. H. Vogel), 349
Knight, F. B. and R. H. Franzen, Textbook Selection, (R), (Herman Remmers), 447
Kuhlmann, F.: Handbook of Mental Tests, (R), (Laura M. Chassell), 165

Labor, Compulsory School Attendance and Child, (R), Forest Chester Ensign, (C. C. McCracken), 73
Laboratory Plan, The Dalton, (R), Evelyn Dewey, (Harriet O'Shea), 167
Languages, Prognosis Tests of Ability to Learn Foreign, Thos. H. Briggs, 386
Legal Status of the City School Superintendent, The, (R), J. Cayce Morrison, (F. C. Landsittell), 352
Letter-Writing: A Curriculum Study in English Composition, Roy Ivan Johnson, 423
Lincoln, E. A.: The Mental Age of Adults, 133
Literature, Functions of Tests and Scales as Found in Recent Educational, A. G. Capps, 204
Los Angeles, Curriculum-Making, (R), Franklin Bobbitt, (W. W. Charters), 347

Material, Interest Factors in Primary Reading, (R), Fannie W. Dunn, (Frances Jenkins), 69
Mathematics, Connersville Course of Study in Elementary, (R), G. M. Wilson, (Madge Perrill), 262
Measure of Achievement, The Efficiency Quotient as a, T. L. Torgerson, 25
Measuring Scale for Freehand Drawing, A, (R), L. W. Kline and Gertrude L. Carey, (Wm. H. Vogel), 349
Mental Age of Adults, (E), F. N. Freeman, 441
Mental Age of Adults, The, E. A. Lincoln, 133
Mental Ability, How Much does a teacher Need? W. B. Bliss, 33
Mental Traits with Age Determined by Annual Re-tests, Changes in, (R), F. D. Brooks, (V. A. C. Henmon), 72
Mental Growth Curve of Normal and Superior Children, (R), Bird T. Baldwin, (G. M. Ruch), 261
Mental Tests, Handbook of, (R), F. Kuhlmann, (Laura M. Chassell), 165
Minimum Essentials in Place Geography, P. R. Stevenson, 300
Montessori Experiments in a Large Infants School, (R), Mary

Blackburn, (Luella A. Palmer), 267
Morrison, J. Cayce: The Legal Status of the City School Superintendent, (R), (F. C. Landsittell), 352
Myerson, Abraham: The Foundations of Personality, (R), (Ernest N. Henderson), 70

Nationality and School Progress: A Study in Americanization, (R), Riverda Harding Jordan, (Josephine D. Mason), 66
New York State, Rural School Survey of, (R), Julian E. Butterworth, (P. R. Stevenson), 448
Niemeyer, N.: Children and Childhood, (R), (A. Sophie Rogers), 169
Normal and Superior Children, Mental Growth Curve of, (R), Bird T. Baldwin, (G. M. Ruch), 261
North Central States from 1860 to 1918, The Development of High-School Curricula in the, (R), John Elbert Stout, (H. A. Hollister), 162

Oral and Silent Reading, (R), C. R. Stone, (John A. O'Brien), 452
Otis Tests, Comparison of the Binet-Simon and, S. C. Garrison and James S. Tippett, 42

Pearson Correlation Co-Efficient, A Data Sheet for the, L. L. Thurstone, 49
Percentile Ranks, Calculation and Interpretation of, L. L. Thurstone, 225
Personality, The Foundations of, (R), Abraham Myerson, (Ernest N. Henderson), 70
Physical Make-up of a book upon Children's Selection, The Effect of the, (R), Florence Eilau Bamberger, (Harold E. Burtt), 170
Pitfalls in the Administrative Use of Intelligence Tests, some, M. R. Trabue, 1
Place Geography, Minimum Essentials in, P. R. Stevenson, 300
Powers, S. R.: A Comparison of the Achievements of High School, and University Students in Certain Tasks in Chemistry, 332
Primary Reading Material, Interest Factors in, (R), Fannie W. Dunn, (Frances Jenkins), 69
Principal as Executive, The, (E), B. R. Buckingham, 254
Problems, Intelligence Quotients of Adults and Related, W. F. Dearborn, 307
Problems, Teaching Geography by, (R), Ehrlich E. Smith, (Douglas C. Ridgley), 75
Professor Terman's Determinism: A Rejoinder, W. C. Bagley, 371
Prognosis Tests of Ability to Learn Foreign Languages, Thos. H. Briggs, 386
Program of Religious Education, The New, (R), G. H. Betts, (L. A. Weigle), 265
Progress and Research, (E), B. R. Buckingham, 257
Progress, Nationality and School: A Study in Americanization, (R), Riverda Harding Jordan, (Josephine D. Mason), 66
Psychological Determinist; or Democracy and the I.Q., (E), Lewis M. Terman, 57
Psychological and Educational Tests in the Public Schools of Winchester, Virginia, (R), W. F. Dearborn and Alexander Inglis, (W. W. Coxe), 260

Psychology of Arithmetic, The (R), Edward L. Thorndike, (Cyrus D. Mead), 163
Public Schools of Winchester, Virginia, Psychological and Educational Tests in the, (R), W. F. Dearborn and Alexander Inglis, (W. W. Coxe), 260
Public School Survey, Technique of the, J. B. Sears, 281
Pupils, Can Teachers Select Bright and Dull? G. F. Verner, 126
Pupils, Classification by Test of Intelligence of Ten Thousand First-Grade, Chas. S. Berry, 185
Pupils in the Elementary Grades, Theory of Differential Education as applied to Handicapped, J. E. W. Wallin, 209

Quotient as a Measure of Achievement, The Efficiency, T. L. Torgerson, 25
Quotients of Adults and Related Problems, Intelligence, Walter F. Dearborn, 307

Reaction, (E), S. A. Courtis, 159
Reading and Spelling with Special Reference to Disability, A Study of, Arthur R. Gates, 12
Reading Material, Interest Factors in Primary, (R), Fannie W. Dunn, (Frances Jenkins), 69
Reading, Empirical Studies in School, (R), James F. Hosic, (S. S. Brooks), 68
Reading Tests, Recent Developments in Silent, C. R. Stone, 102
Reading, Scientific Determination of the Content of the Elementary School Course in, (R), W. L. Uhl, (C. R. Stone), 499
Reading, Silent and Oral, (R), C. R. Stone, (John A. O'Brien), 451
Recent Developments in Silent Reading Tests, C. R. Stone, 102
Recent Educational Literature, Functions of Tests and Scales as Found in, A. G. Capps, 204
Related Problems, Intelligence Quotients of Adults and, W. F. Dearborn, 307
Retrenchment, (E), B. R. Buckingham, 63
Research and Progress, (E), B. R. Buckingham, 257
Responses, Criticism of Tests Requiring Alternate, H. H. Hahn, 236
Retail, How to Sell at, W. W. Charters, (R), (B. R. Buckingham), 451
Revision of the Binet-Simon Tests, Herring, (R), John P. Herring, (P. R. Stevenson), 259
Rhodes, Chas. E.: Effective Expression, (R), (W. L. Graves), 351
Richardson, J. W.: Another Educational Campaign, 97
Rugg-Clark Tests in Algebra, Comparative Validity of the Hotz Scales and the, Eleanore Harris and Frederick S. Breed, 393
Rural School Survey of New York State, (R), Julian E. Butterworth, (P. R. Stevenson), 450

Scale for Measuring Freehand Drawing, (R), L. W. Kline and G. L. Carey, (Wm. H. Vogel), 349
Scale, Binet, The Classification of Kindergarten Children for First-Grade by means of the, Chas. D. Dawson, 413
Scales, Comparative Validity of the Hotz and the Rugg-Clark Tests in Algebra, Eleanore Harris and Frederick S. Breed, 393

Scales as Found in Recent Educational Literature, Functions of Tests and, A. G. Capps, 204
School Finances, The Control of City, (R), Geo. W. Frasier, (B. R. Buckingham), 259
School, Montessori Experiments in a Large Infants, (R), Mary Blackburn, (Luella A. Palmer), 267
School Survey, Technique of the Public, J. B. Sears, 281
School Superintendent, The Legal Status of the City, (R), J. Cayce Morrison, (F. C. Landsittell), 352
School Progress, Nationality and: A Study in Americanization, (R), Riverda H. Jordan, (Josephine D. Mason), 66
School Reading, Empirical Studies in, (R), James F. Hosic, (S. S. Brooks), 68
School Attendance and Child Labor, Compulsory, (R), Forest Chester Ensign, (C. C. McCracken), 73
School Supervision, Some Essentials in, E. J. Ashbaugh, 116
School Survey of New York State, Rural (R), Julian E. Butterworth, (P. R. Stevenson), 488
Schools of Winchester, Virginia, Psychological and Educational Tests in the Public, (R), W. F. Dearborn and Alexander Inglis, (W. W. Coxe), 260
Scientific Determination of the Content of the Elementary School Course in Reading, (R), W. L. Uhl, (C. R. Stone), 449
Sears, J. B.: Technique of the Public School Survey, 281
Selection, The Effect of the Physical Make-up of a book upon Children's, (R), Florence Eilau Bamberger (Harold E. Burtt), 170
Selection, Textbook, (R), R. H. Franzen and F. B. Knight, (Herman Remmers), 445
Sell and Retail, How to, (R), W. W. Charters, (B. R. Buckingham), 450
Seniors, The Intelligence of High School, (R), Wm. F. Book, (Dorothy H. Yates), 447
Silent and Oral Reading, (R), C. R. Stone, (John A. O'Brien), 452
Silent Reading Tests, Recent Developments in, C. R. Stone, 102
Simon-Binet Tests, Herring Revision of the, (R), John P. Herring, (P. R. Stevenson), 259
Simon-Binet and Otis Tests, Comparison of the, S. C. Garrison and James S. Tippett, 42
Skill in Teaching, Acquiring, (R), James R. Grant, (Dwight H. Stevenson), 448
Smith, Ehrlich E.: Teaching Geography by Problems, (R), (Douglas C. Ridgley), 75
Snedden, David: Civic Education, (R), (Chas. C. Peters), 350
Some Essentials in School Supervision, E. J. Ashbaugh, 116
Some Pitfalls in the Administrative Use of Intelligence Tests, M. R. Trabue, 1
Special Reference to Disability, Study of Reading and Spelling with, Arthur R. Gates, 12
Spelling with Special Reference to Disability, Study of Reading and, Arthur R. Gates, 12
Status of the City School Superintendent, The Legal, (R), J. Cayce Morrison, (F. C. Landsittell), 352
Stecher, Lorle I. and Bird T. Bald-

win: Mental Growth Curve of Normal and Superior Children, (R), (G. M. Ruch), 261
Stevenson, P. R.: Minimum Essentials in Place Geography, 300
Stone, C. R.: Silent and Oral Reading, (R), (John A. O'Brien, 452
Stone, C. R.: Recent Developments in Silent Reading Tests, 102
Stout, John Elbert: The Development of High-School Curricula in the North Central States from 1860 to 1918, (R), (H. A. Hollister), 162
Students in Certain Tasks in Chemistry, A Comparison of the Achievements of High School and University, S. R. Powers, 332
Study, Training for Effective, (R), F. W. Thomas, (B. H. Bode), 348
Study in Elementary Mathematics, Connersville Course of, (R), G. M. Wilson, (Madge Perrill), 262
Study in English Composition: Letter-Writing, A Curriculum, Roy Ivan Johnson, 424
Study of Reading and Spelling with Special Reference to Disability, Arthur R. Gates, 12
Studies in School Reading, Empirical, (R), James F. Hosic, (S. S. Brooks), 68
Survey of New York State, Rural School, (R), Julian E. Butterworth, (P. R. Stevenson), 488
Survey, Technique of the Public School, J. B. Sears, 281
Superintendent, The Legal Status of the City School, (R), J. Cayce Morrison, (F. C. Landsittell), 352
Superior Children, Mental Growth Curve of Normal and, (R), Lorle I. Stecher and Bird T. Baldwin, (G. M. Ruch), 261
Supervision, Some Essentials in School, E. J. Ashbaugh, 116

Tasks in Chemistry, A Comparison of the Achievements of High School and University Students in Certain, S. R. Powers, 332
Test of Intelligence of Ten Thousand First-Grade Pupils, Classification by, Chas. S. Berry, 185
Test in the Public Schools of Winchester, Virginia, Psychological and Educational, (R), W. F. Dearborn and Alexander Inglis, (W. W. Coxe), 260
Test Validation, Attempts at, Raymond Franzen, 145
Tests, Some Pitfalls in the Administrative use of Intelligence, M. R. Trabue, 1
Tests, Comparison of the Binet-Simon and Otis, S. C. Garrison and James S. Tippett, 42
Tests, in Education, Trade, (R), Herbert Anderson Toops, (J. Cayce Morrison), 74
Tests, Recent Developments in Silent Reading, C. R. Stone, 102
Tests and Scales as Found in Recent Educational Literature, Functions of, A. G. Capps, 204
Tests Requiring Alternate Responses, Criticism of, H. H. Hahn, 236
Tests, Herring Revision of the Binet-Simon, (R), John P. Herring, (P. R. Stevenson), 259
Tests of Ability to Learn Foreign Languages, Prognosis, Thos. H. Briggs, 386
Tests in Algebra, Comparative Validity of the Hotz Scales and the Rugg-Clark, Eleanore Harris and Frederick S. Breed, 393

Teacher, need? How Much Mental Ability does a, W. B. Bliss, 33
Teachers, Select Bright and Dull Pupils, Can, G. F. Verner, 126
Teaching, Acquiring of Skill in, (R), James R. Grant, (Dwight H. Stevenson), 446
Teaching to Think, (R), Julius Boraas, (N. S. Maddox), 265
Technique of the Public School Survey, J. B. Sears, 281
Terman, Lewis M.: The Psychological Determinist; or Democracy and the I.Q., (E), 57
Terman's Determinism: A Rejoinder, W. C. Bagley, 371
Textbook, Selection, (R), R. H. Franzen and F. B. Knight, (Herman Remmers), 447
Textbooks, The Difficulty of, (E), B. R. Buckingham, 444
Tippett, James S. and S. C. Garrison: Comparison of the Binet-Simon and Otis Tests, 42
Toops, H. A.: Determining Chronological Age in Decimal Parts of a Year, 438
Toops, H. Anderson: Trade Tests in Education, (R), (J. Cayce Morrison), 74
Torgerson, T. L.: The Efficiency Quotient as a Measure of Achievement, 25
Thomas, F. W.: Training for Effective Study, (R), (B. H. Bode), 348
Thorndike, Edward L.: The Psychology of Arithmetic, (R), (Cyrus D. Mead), 163
Thurstone, L. L.: A Data Sheet for the Pearson Correlation Co-Efficient, 49
Thurstone, L. L.: Calculation and Interpretation of Percentile Ranks, 225
Trabue, M. R.: Some Pitfalls in the Administrative Use of Intelligence Tests, 1
Training for Effective Study, (R), F. W. Thomas, (B. H. Bode), 348
Traits with Age Determined by Annual Re-tests, Changes in Mental, (R), Fowler Dell Brooks, (V. A. C. Henmon), 72

Uhl, W. L.: Scientific Determination of the Content of the Elementary School Course in Reading, (R), (C. R. Stone), 449
United States History, An American Citizenship Course in, (R), Fannie Fern Andrews, (R. M. Tryon), 67
University Students in Certain Tasks in Chemistry, A Comparison of the Achievements of High School and, S. R. Powers, 332
Use of Intelligence Tests, Some Pitfalls in the Administrative, M. R. Trabue, 1

Validation, Attempts at Test, Raymond Franzen, 145
Validity of the Hotz Scales and the Rugg-Clark Tests in Algebra, Comparative, Eleanore Harris and Frederick S. Breed, 393
Verner, G. F.: Can Teachers Select Bright and Dull Pupils? 126
Virginia, Winchester, Psychological and Educational Tests in the Public Schools of, W. F. Dearborn and Alexander Inglis, (R), (W. W. Coxe), 260
Voelker, Paul Frederick, The Function of Ideals and Attitudes in Social Education, (R), (W. J. Osburn), 168

Wallin, J. E. W.: Theory of Differential Education as Applied to Handicapped Pupils in the Elementary Grades, 209
Winchester, Virginia, Psychologi-

cal and Educational Tests in the Public Schools of, (R), W. F. Dearborn and Alexander Inglis, (W. W. Coxe), 260
Wilson, G. M.: Connersville Course of Study in Elementary Mathematics, (R), (Madge Perrill), 262
Writing-Letter: A Curriculum Study in English Composition, Roy Ivan Johnson, 424
Wood, Will C.: Adjustments in California, 326

JOURNAL *of* EDUCATIONAL RESEARCH

Volume VI JUNE, 1922 Number 1

SOME PITFALLS IN THE ADMINISTRATIVE USE OF INTELLIGENCE TESTS[1]

M. R. TRABUE
Teachers College, Columbia University

As long as intelligence tests were used only by the psychologists who originated them, the public had little cause for alarm. In the laboratory and in the hands of trained specialists, mental examinations were not dangerous. Their significance and their limitations were appreciated, and such action as was based on them was reasonable and conservative. The possibility of measuring such qualities as mental capacity and intellectual alertness has aroused the public interest, however, and intelligence tests are now being applied in all parts of the country and by persons with little or no preparation for such work.

It is said to be extremely dangerous in certain parts of the world to walk through the open country at night, on account of the deserted mine-shafts and unprotected pits that await the feet of one who travels without a light. In order to insure one's safety in such a situation, there would be three main possibilities: one might secure an adequate lantern to light his path; he might sit down patiently and wait for daylight before going further; or he might move forward slowly and cautiously, feeling out each step in advance of taking it.

The dangers that beset the path of would-be mental examiners are of much the same nature. There are pitfalls on every hand. Unless his eyes have been opened to these dangers by thorough training and extensive experience, one should not attempt to give intelligence tests. If the attempt is made, it should be made with extreme caution. It would be much better to wait for expert guidance in this field than to plunge blindly ahead and thereby involve scores of innocent boys and girls in an educational and vocational snare, which would affect their lives profoundly and in many cases permanently.

---

[1] Address delivered before the joint meeting of the Department of Superintendence and the National Society for the Study of Education held at Chicago, February 28, 1922.

One of the sources from which school officers have drawn their ambition to use intelligence tests in administrative work, has been the success of such tests in classifying hundreds of thousands of recruits in the United States Army. I was closely connected with the army experiment. The success of the tests in military camps was a relative matter. The mistakes that the psychological examiners might make were less serious than those that were certain to be made by the unsupported guesses of military officers, who for the most part were wholly unacquainted with their men.

In the army we were classifying adult men for temporary service, but in the schools we are classifying immature boys and girls, with the frequent assumption that the assignments made will hold for the entire period of their school courses, if not for life. The work in the army camps was carried on, or at least carefully supervised, by trained psychologists. Much of the work in schools, however, is being done by teachers and principals who have almost no special training or experience in this field. It is time, therefore, for us to make a careful examination of the premises and to discover the pitfalls that beset this relatively uncharted path.

PITFALLS IN ONE'S UNDERSTANDING OF THE NATURE OF THE TESTS

An inadequate understanding of the scope and significance of the tests may lead to serious difficulties. "Intelligence" is a big word, and it signifies to the average person a broad range of abilities. Each of the group intelligence scales at present available involves only a few of these abilities, and no two of these scales measure exactly the same combination of qualities. We know very little about the specific traits that should be included in a comprehensive test of *general intelligence,* and still less about the relative weight that should be attached to the different tests to be included. We are certain that we do not yet have any system of tests in which exactly the proper elements are combined in exactly the proper proportions to make a perfect test of general intelligence. What then do the intelligence tests actually measure? What are the criteria by which we are judging the intelligence of pupils?

As teachers, we tend to think of intelligence in the rather narrow terms of school success. The child who prepares his lessons easily and well is "intelligent." The one who has difficulty in understanding words, symbols, and complex ideas is "rather stupid." Practically all of the present intelligence tests are built on the assumption that success in school work is only another measure of intelligence. The pupils who make high scores on the intelligence tests are the ones who enjoy and can easily master the books and abstract ideas offered in the schools.

Miss Gambrill has recently demonstrated that successful business men do not usually come from the group of honor students at college. The men who win Phi Beta Kappa keys and high grades in their college studies do not win the largest incomes after graduation. Those who win the large incomes after graduation tend to come from the students who make only average grades in college work. The honor students in college tend to become teachers. In other words, the ability to make a favorable impression on college instructors, through brilliant recitations and excellent examination papers, is exactly the same ability as the instructors themselves possess. In assigning grades for success in school or college work, one tends to rate very high those whose abilities are like his own and to rate lower those who are mentally less like him. Is it right, however, for us to go still further and say that those who are unlike us in their abilities are less intelligent?

I know of a woman, forty years of age, whose rating on the intelligence tests is about equivalent to that of a fourth-grade school pupil. In actual fact she gave up school work when she had reached the fifth grade, because she did not care for the things she was learning. She is now a successful American business woman, in spite of having been left alone in the world with no money and with an I. Q. of not more than 75. At the present time she is the owner and manager of a small factory. For several years she was the housekeeper at one of the finest Fifth Avenue hotels, where she successfully directed the work of a corps of approximately fifty maids, three carpenters, two decorators and a plumber. In spite of the evidence of the tests, I insist that she is intelligent.

It is certainly a mistake to claim that the group tests used in our schools are measures of *general* intelligence. After all, is it general intelligence that we wish to measure? Is it not really more important for our purposes to measure the specific type of intelligence which enables children to succeed in the kind of work offered by the schools, than to measure intelligence in general? A test of general health, which gave no more specific information than that the patient was "very sick," would not be any more useful in medical practice than a general intelligence test would be in educational practice. The good physician seeks for specific symptoms rather than for general conditions. As I see it, the school administrator would do well to acknowledge frankly, when he uses a group intelligence test, that he is measuring certain elements of a relatively specific ability to do abstract thinking with words and symbols. Such an acknowledgment would go far toward clarifying the issues raised by the results of tests.

In view of the specific nature of the intelligence tests, it is unfortunate that the custom has grown up of speaking of one child as "superior" to another, when the only evidence of such superiority is that he has made a larger score on a given test. The evil of this custom is more clearly seen when one child is reported as "inferior" to another. How do we know that the perfect all-round intelligence test might not weight social intelligence, or mechanical intelligence, or some other type of ability, much more heavily than the academic type we now can measure? Do we know that the child making a low score in abstract intelligence might not be rated much higher if measured by a truly general intelligence test? The words "superior" and "inferior" give far too great an importance to the type of tests that are now available. College graduates in business make a higher score, on the salary scale, than college graduates in education, but we do not for that reason need to speak of one of these groups as superior to the other.

There is one further danger of misunderstanding with regard to the character of intelligence tests. Studies of their correlations have proved that they predict success in school more accurately than it is predicted by the judgment of the average teacher. As a consequence of this, there has been a tendency to say that the results of intelligence tests should therefore be used as the basis for the promotion and classification of pupils, more or less regardless of the judgments of teachers. This tendency must be corrected. Intelligence test scores should never be substituted for common sense and good judgment. The tests should supplement the teacher's judgment but should not take its place. Where common sense and reason seem to oppose the evidence of the tests with regard to a pupil, the tests should not be accepted as conclusive. The time may come when an intelligence test score may be trusted, even though it contradicts the thoughtful judgments of teachers, but that time is not yet.

Pitfalls of Exaggerated Confidence in the Accuracy of the Tests

The very language in which the results of tests are discussed affords numerous pitfalls, since it tends to create an unwarranted confidence in the accuracy of the instruments used. The word "measurements," which has come to be applied to the results of almost any test, suggests a high degree of accuracy which cannot always be justified. When one goes still further, however, and claims that his test results are "scientific measurements," there are few indeed who dare to suggest the extent to which other equally scientific measurements might disagree, especially as regards the abilities of individual pupils.

This suggestion of extreme accuracy in test results is further exaggerated by the numerical character of the scores obtained. A pupil is reported, for example, as having a score of 63 on the National Intelligence Test, while 72.4 is reported as the average score of his group. The difference between this pupil's score and his group average may then be calculated and further elaborate analysis made, as though 63 were an absolutely accurate point, which belonged to the pupil as intimately and constantly as his name or the number of inches in his height.

The mental examiner for a private school in New York City informed me last week that she was afraid to give a second intelligence test to her pupils, for fear it would give different results from the one given last September. Said she, "We classified all our pupils on the basis of their scores in the Otis Tests, and I am afraid the patrons would lose their faith in the accuracy of intelligence tests if a second trial should disagree with the first." Such a remark well illustrates the unwarranted confidence that has been built up, all too generally, in the accuracy of the tests. The sooner this over-confidence is broken up, the better for the future of intelligence testing and intelligence testers.

Another source of false confidence has been the adoption, by the makers of group tests, of the Intelligence Quotient technique, devised by Terman for the Stanford Binet Examination. The truth of the doctrine of the constancy of the I. Q. has not been sufficiently demonstrated in the case of the Stanford Examination itself, but as soon as its truth is asserted in connection with the far less accurate group examinations, a sad state of affairs has been reached in the science of mental measurements. In hundreds of cases, however, pupils have been tested by a group intelligence test, the accuracy of which even the examiner himself would question, after which ratios and quotients have been worked out to the second or third decimal point. All too frequently these derived measures have been discussed and used in administrative reclassifications, as if their accuracy were open to no question or doubt whatever.

Examples of this over-confidence in the accuracy and finality of intelligence tests can be found wherever such tests have been used. In one of the best private schools in New York City, a third-grade pupil who was leading her class and whom the teacher recommended for special promotion was not allowed such a promotion, because it was learned, forsooth, that as a kindergarten child her I. Q. on the Stanford Examination had been found to be only 89. If this I. Q. had been obtained by using one of the various group tests of intelligence, the administrative action would have been still more unwarranted. Unless one is willing to re-examine and to reclassify pupils whenever new and more reliable evidences of ability have been discovered, he might better leave the intelligence tests strictly alone.

A word should be said at this point about the use of the phrase "Intelligence Quotient" or "I. Q." If this stands alone and unidentified, one naturally assumes that it refers to the Stanford Revision of the Binet Scale. A quotient obtained from the Otis Tests, for example, should be clearly identified as such. Likewise with "Mental Age" scores, it is naturally assumed that a pupil reported as having a given mental age was examined by the Stanford Binet Scale. If the age score was obtained from the National Scale A, it should be clearly identified as a "National Scale A Mental Age." Unless all such derived scores are thus identified, as "Miller Mental Ages," "Dearborn Intelligence Quotients," and "Otis I. Q.'s," it will be impossible for other persons than the original examiner to form a correct idea of the significance and accuracy of such measures.

Pitfalls in Reasoning from the Results of Tests

A third type of error is due to loose reasoning about the results of tests. False inferences are too often drawn from the scores obtained. It is quite common, for example, to assume that pupils who make low scores in a so-called intelligence test would succeed relatively better along trade lines, in manual arts, or in unskilled labor. Perhaps they would, but the low scores in the abstract tests do not prove it. The absence of one type of ability does not insure the presence of another. Until adequate tests for the purpose have been applied, the presence of a capacity for manual skill, or the possession of any other type of ability, can be proved at the present time only by an actual try-out of the pupil in the field concerned. Any assignment of a pupil to such "practical studies" on the basis of demonstrated inability to deal with abstract symbols and ideas should be only tentative. Permanent assignment to a course should not be made until after ability in that particular line has been proved.

Another false conclusion, drawn from the results of the so-called intelligence tests, is that pupils who make scores of a certain size will not profit by further training. The tests used are measures of a specific type of ability,—the type which enables pupils to succeed at the academic work offered in the usual public school. If a child makes a low score in the tests, it does not follow that it is useless to train him. The child may indeed have reached his maximum level in abstract work with symbols and words, but there remains an endless number of avenues through which he may broaden his experience and enrich his store of knowledge, even at his relatively low level of abstract thought. There may be other fields of training, also, in which such a child may demonstrate a relatively higher level of ability. It is certainly unfair to refuse further training to a child because the school has in the past failed to recognize the needs of pupils of his type.

At the other extreme there is a pitfall of just as serious a nature,—the assumption that those who make high scores on the so-called intelligence tests do not need special attention or training, because they will take care of themselves. It is true that pupils with great ability as shown by these tests are usually able to master the ordinary school subjects with little effort and great speed, but unless they are stimulated and directed they do not develop as they should along the lines for which their special abilities fit them. Stenquist has recently found a negative correlation of 0.40 between effort and intelligence, both qualities in several hundred pupils being measured as accurately and objectively as possible.

The pupils with high abstract intelligence test scores learn what is offered for the average pupils and then loaf. Their need is not merely for more work but also for a different type of training. The average citizen needs to be taught how to verify the information supplied by books, magazines, newspapers, and research workers, and how to judge and select the ideas by which the democracy he controls is to be guided. But the ideas from which the average citizen is to choose, and the researches by which information is to be obtained, will be produced by the few who enjoy and are successful in dealing with abstract matters. These few may be detected early in life by means of the present abstract intelligence tests. The present tests will not, however, certify potential leadership in social, political, or industrial life. Other tests will have to be devised for these purposes. Each pupil, whatever may be the character and degree of his natural gifts, must be given training of such a nature as will give the utmost possible return to the community from his particular intellectual endowment.

Pitfalls in the Treatment Applied to Conditions Discovered by Tests

The most serious pitfalls of all are those concerned with the treatment applied to situations revealed by intelligence tests. As long as one does nothing but talk or think about the results of such tests, he is not likely to do a great deal of harm, unless someone else hears and acts upon his thoughts.

First of all, an administrative officer in a school system should recognize that any re-organization of classes or reclassification of pupils which he may make with the aid of intelligence tests is not in itself a desirable educational objective. When pupils have been sorted into homogeneous groups, they have not thereby been improved. Tests of ability along various lines, and the grouping together of pupils of like abilities, furnish a starting point and an opportunity for an educational program that will lead pupils to success and happiness rather than to failure and discouragement; but they do not in themselves provide such a program. All of the intelligence, patience, and skill of a competent corps of supervisors will be needed to develop and to carry on such an enterprise. It will not run automatically.

As soon as one begins to make administrative use of the intelligence tests, teachers and supervisors must be trained to make proper adjustments of their instruction to the characters and abilities of their groups. This is not an easy task. There will be a strong tendency for teachers to slip back into the more traditional ways of dealing with their classes. Supervisors will need to remind the teachers constantly of the special opportunities they wish to provide for the pupils so unusually well-examined and classified.

In many cases teachers will have to alter not only their methods but their social attitudes and their conceptions of democracy. Instead of trying to turn the eyes of all their pupils toward academic work or the learned professions as the ultimate goal in life, they must recognize and teach that effective service to the community is the highest ideal. It must be clear to them that it is just as noble and worthy to be an efficient butcher or teamster as to be a good teacher, or indeed, that an effective street-sweeper may actually be worthy of more public honor and respect than a successful lawyer.

Above all, teachers must cease thinking and speaking of the "higher professions" when referring to types of work similar to their own. The only sense in which one man's work is "higher" than another's is that, considering his talents, the one is contributing more effectively than the other to the common welfare. Teachers must recognize and train other types of ability than those measured by the present abstract intelligence tests, and there must not be any suggestion that those whose abilities are different from their own are of a "lower order." The early discovery of the type of ability possessed by a pupil must be recognized as making possible for him a freedom from failure in lines not suited to him, and as providing him with greater rather than smaller opportunities for *effective* training.

Each pupil must come to think of failure to measure up to his own possibilities as the great defect. He must honor every other person who, for the good of the community, is making the most of his talents, regardless of the nature of those talents. Those few pupils whose inheritance makes it possible to train them to think original thoughts are not, on that account, entitled to any greater respect or to any less attention than the more numerous pupils who must be trained to judge and select carefully the thoughts and ideals upon which to base the activities of the social organization to which all belong and for which all are laboring. The inferior person must be recognized by all as the one who does less than his own best.

Such a program requires time. It also requires an enormous amount of sound common sense, a broad social philosophy, a thorough training in psychology and scientific method, and an infinite patience. For the present, and until the science of mental measurements is much older, more highly developed, and better understood, the psychologists and the school administrators must cooperate at every step. Each needs the constant assistance of the other, if the pitfalls that surround the administrative use of intelligence tests are to be avoided. For the sake of the pupils and for the safety of democracy, psychologists and administrators must continue to work hand in hand. "United we stand, divided we fall."

A STUDY OF READING AND SPELLING WITH SPECIAL REFERENCE TO DISABILITY[1]

Arthur I. Gates
Teachers College, Columbia University

The purposes of this investigation were as follows: (1) to devise, as well as we could, a technique for the diagnosis of disability in reading and spelling; (2) to discover the constituent factors involved in acquiring these functions by a study of both good and poor readers; (3) to ascertain the causes of such disabilities as were found; and (4) to try out certain remedial forms of instruction.

For this purpose approximately 135 cases of whom 25 or 30 may be considered as "seriously" retarded in reading or spelling, or both have been studied in detail. Very few of the subjects were of less than average intelligence. (See Table I).

TESTS AND EXAMINATIONS GIVEN

I. Tests of General Mental Ability.
   1. Stanford-Binet, most subjects.
   2. Nine group tests, all subjects.

II. Tests of Achievement in Reading, Writing, Arithmetic, and Spelling.

III. Tests of several forms of Visual Perception.
   1. Ability to detect small differences between pairs of printed materials: four tests, employing as materials, pictures, digits, nonsense material, and words.
   2. Ability to detect one printed form among several others: four tests, geometrical figures, A.'s, K. I.'s, and digits.
   3. Ability to recognize the correct form of a word among incorrect forms—recognition spelling.
   4. Proof-reading test.

IV. Ability to pronounce single words (pronunciation test devised by the writer) and to read orally (Gray's Oral Reading Test).

V. Rate and Depth of Comprehension of words and sentences.
   1. Holley's Visual Vocabulary Test and a similar test arranged by the writer.

---

[1] A paper read at the meeting of the National Association of Directors of Educational Research in Chicago, March 1, 1922. The study is a contribution of the Department of Educational Research of the Scarborough School, Scarborough, New York.

  2. The Burgess, Brown, Courtis, Monroe, Thorndike-McCall Reading Tests and the Woodworth-Wells Direction Test.

VI. Spelling Tests.
  1. A select list of 36 words for oral spelling, given individually.
  2. A select list of 36 words for written spelling, given individually.
  3. Sentences for dictation; written spelling.
  4. A composite of spelling ability comprising 186 words from various columns of the Ayres-Buckingham Scale.

VII. Understanding of Spoken Language.
  1. Terman's Oral Vocabulary Test.
  2. Oral sentences and directions.

VIII. Tests of the Sensory and Motor Mechanisms.
  1. Tests of vision and hearing.
  2. Tests of general motor control, writing ability, and articulation.
  3. Eye-movements in reading.
  4. Lip-movements, etc., in reading.

IX. Emotional, Temperamental, and Other Traits.
  1. Inertia, laziness, disinclination.
  2. Instability.
  3. Poor health, etc.

The following methods of analysis were employed: (1) analysis of group relations by means a simple, partial, and multiple correlations;[2] (2) a comparison of the achievements of the backward cases with those of the normal cases in the several tests; and (3) an analytical study of the responses of backward cases in certain individual examinations.

Limitations permit only the briefest summary of some of the outstanding conclusions, in most instances without a presentation of the substantiating evidence which will be published in detail later.[3]

---

[2] The Pearson products moment formula for simple correlations and Kelley's procedure for partial and multiple correlations were used. Kelley, T. L., *Tables to facilitate the calculation of partial coefficients of correlation and regression equations*, Stanford University Publications in Education, Vol. I, No. I, 1921.

[3] In a forthcoming monograph in the *Teachers College Contributions to Education*.

Defects Associated with Disability in Reading

DEFECTS OR DEFICIENCIES OF THE CENTRAL NERVOUS SYSTEM

General Mental Ability—That the relation between general mental ability and these verbal functions is close, is well known. Terman, Burt, Tredgold, and others find that it is quite futile to attempt to teach children or adults with I. Q.'s of 0.60 or less to read; in fact, an I. Q. of 0.70 may be considered a minmum essential.[4]

In the present investigation no attempt was made to ascertain the minimum intelligence for reading; we sought rather to eliminate this factor as a cause of disability. The following distributions (Table I) shows the I. Q.'s of those measured by the Stanford-Binet Test. The average is about 116. Those not measured by the Stanford Test had at least average intelligence as indicated by group tests.

TABLE I.—DISTRIBUTION OF I. Q.'S AMONG SUBJECTS TESTED

Stanford I. Q.	80–89	90–	100–	110–	120–	130–	140–	150–	160 and Over	Total
Number of Cases.......	2	8	21	22	19	12	2	4	1	91

TABLE II.—CORRELATIONS OF COMPOSITE OF READING ABILITY WITH INTELLIGENCE

Grade	Stanford Mental Age	Verbal Group Intelligence	Non-Verbal Group Intelligence
III	0.30	0.70	0.15
IV	0.35	0.60	0.24
V	0.58	0.65	0.28
VI	0.71	0.77	0.26
VII	....	0.75	0.17
VIII	....	0.80	0.05
Mean........	0.49	0.71	0.20
M. D........	0.16	0.06	0.07

---

[4] Terman, L. M. *The measurement of intelligence.* pp. 266ff. Burt, Cyril. *Mental and scholastic tests.* London, 1921.

TABLE III.—CORRELATIONS OF COMPOSITE OF SPELLING ABILITY WITH INTELLIGENCE

Grade	Stanford Mental Age	Verbal Group Intelligence	Non-Verbal Group Intelligence
III	0.32	0.56	0.13
IV	0.11	0.33	0.18
V	0.37	0.26	0.21
VI	0.45	0.43	0.18
VII	....	0.44	0.12
VIII	....	0.50	0.08
Mean........	0.31	0.42	0.15
M. D........	0.11	0.08	0.04

TABLE IV.—AVERAGES OF THE CORRELATIONS FOR GRADES IV TO VIII, INCLUSIVE

Tests	Detecting Small Differences Between Pairs of Tests				Cancellation Tests			
	V-1 Pictures	V-2 Digits	V-3 Nonsense	V-4 Words	V-5 Figures	V-6 A's	V-7 K. I.	V-8 Digits
V-1 Pictures........	...	.04	.08	.06	.26	.13	.12	.04
V-2 Digits..........	.04	...	.27	.20	.19	.24	.23	.43
V-3 Nonsense.......	.08	.27	...	.46	.23	.21	.23	.24
V-4 Words..........	.06	.20	.46	...	.10	.14	.05	.11
V-5 Figures.........	.26	.19	.23	.10	...	.32	? .24	.39
V-6 A's.............	.13	.24	.21	.14	.32	...	.58	.46
V-7 K. I...........	.12	.23	.23	.05	.25	.58	...	.35
V-8 2, 3,...........	.04	.43	.24	.11	.39	.46	.35	...
Mean............	.10	.23	.25	.16	.25	.30	.26	.30

Even when these superior children are taken in grade groups, in which the range is very narrow, the correlations of general mental ability of the verbal type with reading and spelling are substantial. In Tables II and III the variables are: (1) Stanford-Binet, requiring about an hour of time; (2) a selection of very *verbal* exercises from group tests listed above to require an hour; and (3) a selection of very *non-verbal* exercises to take an hour. The correlations are similar to those usually found.

It is, however, a notable fact that children of superior intelligence, some of them having I. Q.'s of 120 or more, should have experienced extreme difficulty in learning to read.

Defects or Deficiency in Visual Perception.—Ten tests of visual perception (see division III in the list above) were used. The correlations shown in Table IV are low except where the materials were the same or similar. The implication is that there is no 'general perceptual ability' but many relatively specific perceptual abilities; as many, in fact, as there are classes of items, namely, letters, words, digits, geometrical figures, etc.

Detailed study, of both good and poor readers and spellers who do poorly in all perceptual tests discloses the fact that inadequate vision, eye muscle control, sluggishness of motor reaction, inability to attend persistently, etc., rather than any defect of 'general perceptual ability' are causal. When these factors are eliminated, the non-readers and non-spellers were quite as good as the others in perception of all items except words. The abilities to perceive digits, figures, etc., show very low correlations with reading and spelling.

The several tests involving words show high intercorrelations as shown in Table V. They are also substantially correlated with reading and spelling. By the technique of partial and multiple

TABLE V.—CORRELATIONS[*] OF SPELLING AND READING WITH OTHER VARIABLES

	WITH SPELLING	WITH READING	RECOGNITION SPELLING	PROOF READING
Test V-1, Small differences in drawings...	–0.04	0.05	0.01	0.15
Test V-2, Small differences in digits......	–0.05	0.14	0.01	0.09
Test V-5, Cancellation, geometrical figures	0.14	0.16	0.13	0.17
Test V-6, Cancellation, A's.............	0.17	0.14	0.13	0.16
Test V-7, Cancellation, K. I............	0.04	0.04	0.17	0.09
Test V-8, Cancellation, 2, 3,............	–0.03	0.11	0.08	0.11
Test V-3, Small differences, nonsense.....	0.35	0.27	0.32	0.35
Test V-4, Small differences, words.......	0.55	0.39	0.47	0.56

[*] r is the mean of the r's for Grades III to VIII.

TABLE VI. SIMPLE, PARTIAL, AND MULTIPLE CORRELATIONS BETWEEN THE VARIABLES (1) READING, (2) RECOGNITION SPELLING, (3) PROOF READING, AND (4) ABILITY TO DETECT SMALL DIFFERENCES BETWEEN WORDS

SIMPLE CORRELATIONS	PARTIAL CORRELATIONS FIRST ORDER	PARTIAL CORRELATIONS SECOND ORDER
$r_{12} = 0.47$	$r_{12.3} = 0.31$	$r_{12.34} = 0.26$
$r_{13} = 0.45$	$r_{13.2} = 0.26$	$r_{13.24} = 0.19$
$r_{14} = 0.39$	$r_{13.4} = 0.30$	$r_{14.23} = 0.12$
$r_{23} = 0.53$	$r_{14.2} = 0.22$	
$r_{24} = 0.47$	$r_{14.3} = 0.19$	
$r_{34} = 0.56$	$r_{23.4} = 0.37$	
	$r_{24.3} = 0.24$	
	$r_{34.2} = 0.42$	

MULTIPLE CORRELATIONS

$R_{1.23} = 0.524$

$R_{1.234} = 0.54$

correlations (see Tables VI and VII) a factor appears, which is common to these functions and sufficiently important to give substantial correlations, other things being eliminated. By hypothesis, this common factor is defined as an ability to perceive clearly the significant details of words. This hypothesis was verified by a series of analytical studies of poor readers.

The multiple correlations of the perception tests with spelling (Table VII) are higher than those with reading. This suggests the possibility that those subjects who possess a

TABLE VII.—SIMPLE, PARTIAL, AND MULTIPLE CORRELATIONS BETWEEN THE VARIABLES, (1) SPELLING, (2) RECOGNITION SPELLING, (3) PROOF READING, AND (4) ABILITY TO DETECT SMALL DIFFERENCES BETWEEN PAIRS OF WORDS

SIMPLE CORRELATIONS	PARTIAL CORRELATIONS FIRST ORDER	PARTIAL CORRELATIONS SECOND ORDER
$r_{12} = 0.63$	$r_{12.3} = 0.54$	$r_{12.34} = 0.50$
$r_{13} = 0.39$	$r_{12.4} = 0.50$	$r_{13.24} = -0.08$
$r_{14} = 0.55$	$r_{13.2} = 0.09$	$r_{14.23} = 0.38$
$r_{23} = 0.53$	$r_{13.4} = 0.12$	
$r_{24} = 0.47$	$r_{14.2} = 0.37$	
$r_{34} = 0.56$	$r_{14.3} = 0.44$	

MULTIPLE CORRELATIONS

$R_{1.34} = 0.69$

$R_{1.234} = 0.69$

very favorable form of word perception are, to some extent, learning (or relearning) to spell during the course of ordinary reading. This conclusion is implied, but not proved, by the fact that very poor readers were invariably very poor spellers, although there were, for other reasons, good readers who were not good spellers.

Congenital Defects of Special Cortical Areas.—No evidence was found to substantiate such diagnoses as "congenital word-blindness," "alexia," etc. These concepts have been questioned by several writers.[5] It is usually impossible either to prove or disprove them. If they exist they are very, very rare.

Other concepts such as defects of "visual memory," "auditory memory," "inability to associate auditory and visual symbols," "defects of visual imagery," or defects in "visualization," "special disability," etc., must finally be explained in terms of defect or deficiency of the central nervous system and are not essentially different concepts from "word-blindness." Neither neurologists or psychologists are in agreement even among themselves concerning the validity of these several concepts, and our procedure was to consider them only as a last resort in diagnosis.

DEFECTS OF THE SENSORY MECHANISMS

Serious defects of vision were discovered in four cases. Milder defects were found in three others, and in many cases such defects probably existed uncorrected at the period of learning to read. Serious visual defects are without doubt a complete or at least contributory cause of disability in reading and spelling. The defects of hearing were not so common. One case of inability to discriminate tones differing by less than 35 vibrations per second was found. It is doubtful whether this contributed greatly to difficulty in reading or spelling.

DEFECTS OF MOTOR MECHANISMS

Marked motor sluggishness and incoordination were shown in three cases and to a less pronounced degree in two others.

---

[5] Hollingworth, L. S., *Psychology of special disability in spelling.* Teachers College Contributions to Education, 1917. Freeman, F. N., "Clinical study as a method in experimental education." *Journal of Applied Psychology,* March, 1920.

These five and several additional cases were found to be defective in writing ability. Although the approximate correlation between writing and spelling or reading was but 0.15, seriously defective writing does produce a handicap in spelling. In one or two cases a general carelessness was probably responsible for both poor writing and poor spelling.

Defective articulation.—According to our appraisal, defective articulation was definitely associated with inferior reading and spelling in one case. The causal connection was indirect, inasmuch as it appeared that habits of hurrying in order to avoid stammering, when the subject encountered difficult words, had led to a variety of inappropriate habits which inhibited the development of effective methods of attack. A similar solution was suggested but less certain in a second case.

Defective eye-muscular control.—Two cases of defective visual perception grew out of defective coordination of the eye-muscles which was not discovered until strabismus developed. One of these subjects was reported by an oculist as suffering from "defective convergence and word-blindness." Following correction and remedial treatment of the muscular defect, these subjects rapidly learned to read and spell as a result of training in word perception and analysis. The former subject improved much more slowly, due in part, probably, to a continued difficulty in control of eye-movements, and in part, to the persistence of certain unfavorable reading habits.

Inappropriate eye-movements.—In cases showing no organic defect, inappropriate eye-movements usually accompany all types of difficulty in reading. It appeared, in the cases here studied, that the inappropriate eye-habits were a secondary but necessary result of difficulty in reading however caused. We were unable to alleviate the difficulty in any case tried by means of training directed to eye-movements alone. On the other hand, the facility with which eye-movements readjusted themselves as the result of improvement in reading achieved by other forms of training was very great.

Inappropriate Eye-Voice Span.—Like inappropriate eye-movements, a limited eye-voice span (inability to look ahead of the words being silently or audibly articulated) seems to be a symptom or result of other difficulties rather than the cause of them.

UNFAVORABLE BEHAVIOR OF A GENERAL CHARACTER

Disinclination, inertia, etc.—Under this caption a large number of characteristics—sheer laziness, distaste for school or for mental work, emotional disturbance either temporary or chronic, outside interests, home spoiling, etc.—might be considered. Without doubt, some of the backwardness was due to lack of interest and effort, however occasioned. Disinclination was sometimes caused by inability which in turn was due to ineffective methods or native inaptitude. The function was disliked because it was difficult, because little progress attended application, and because of the chagrin of constant failure. When the pupil had acquired effective methods of study and observed that he really could learn, a new and happy interest was the common result. Some of the pupils, for example, began to read books, a thing that had never previously been done because reading was difficult work. Several cases, however, especially in spelling, failed to develop any considerable interest or initiative. To discover the ultimate cause of such disinclination would require a more intimate acquaintance with the life history and heredity of these cases than we were able to secure.

Emotional and Nervous Instability.—That the unstable or neurotic child is frequently retarded in school work, specially in work such as phonetics, spelling, and arithmetic which require persistent drill, has been frequently noted. Two of our cases were of this type. They were hasty and careless, both in learning and in performing during the tests. Their methods of study were neither effective nor consistent; they tired of the work before they were well under way. Add to the neurotic constitution almost any special inadequacy for the work and pronounced backwardness is the usual result.

UNFAVORABLE TRAINING

General psychology teaches us that we learn to perceive words, objects, distance, etc., in much the same manner that we learn to play the piano, typewrite, or sing. In the case of acquisition of skills, we realize the need and value of instruction. It is equally, perhaps more, needed in learning to perceive words for the reason that the appropriate and inappropriate responses can be detected by the learner only with difficulty. Children left to themselves will seldom discover the economical and effective methods of playing a piano or running a typewriter, and we may expect many to develop very ineffective methods of perceiving words in reading and spelling.

In at least one-third of our cases, lack of training or inappropriate training was responsible, at least in part, for the difficulty in reading and spelling. Some of the common difficulties, caused in whole or in part by inappropriate training or lack of training are: (1) Misrecognition and mispronunciation due to habits of perceiving words vaguely as wholes, or of perceiving certain essential parts vaguely, or of perceiving in some other inappropriate way. (2) Inability or ineffective methods in analyzing and attacking unfamiliar words. (3) Habits of neglecting unfamiliar words accompanied by a search through the material in an effort to comprehend the passage on the basis of such words as are familiar. (4) Mispronunciation of familiar words because of the procedure described in (3). (5) Various types of inappropriate eye-movements because of (1), (2), or (3). (6) Narrow eye-voice span i. e., inability to see ahead of material being pronounced aloud or silently, because of (1), (2), or (3). (7) Super-sufficient reaction to details of words, due to improper training, lack of training, or some other cause. These causes result in a variety of errors, in inability to grasp large units, and often in inability, to recognize long or uncommon words. (8) A variety of unsystematic procedures in word recognition. (9) Failure to comprehend because attention is absorbed in recognition or pronunciation. This occurs among the mechanically good, as well as among the poor oral readers. (10) Over-emphasis of articulatory responses resulting in low speed and various other inadequacies such as expressionless oral reading, rapid slurring reading, tendencies to stammer, nervousness,

TABLE IX.—SCORES IN SPELLING TESTS IN TERMS OF PERCENT OF CORRECT SPELLING OF EQUIVALENT LISTS SELECTED FROM AYRES-BUCKINGHAM SCALE

Grade	Score Nov. 1	Score Jan. 15	Score May 15	Improvement[a] in Years
III	21	40	55	1.7
IV	29	52	69	1.8
V	42	63	82	2.2
VI	40	57	70	2.0
VII[b]	61	63	66	0.4
VIII[b]	59	62	64	0.4

[a] Figures in this column are improvements expressed in terms of average improvements achieved by American schools in general (Ayres norms). Grade III-VI used the special method.

[b] Grades VII and VIII had little specific instruction in spelling.

all cases, both good and bad readers and spellers in grade groups, correlations from 0.50 to 0.90 are found between status at the beginning and at the end of the year. That is to say, those good at the start of a period of training, generally speaking, still excel their initially less gifted companions at the end; and in all likelihood they will still show their superiority ten years hence. While we cannot make all children equally efficient readers and spellers, by means of careful diagnosis and specific remedial instruction, we can make most of them better than they otherwise would be and enable some of them to learn who otherwise might not learn at all.

THE EFFICIENCY QUOTIENT AS A MEASURE OF ACHIEVEMENT

T. L. TORGERSON
Director of Educational Measurements, West Allis, Wisconsin

Users of standard educational tests are no longer content with measuring the results of instruction in the various grades of a school but are demanding an accurate measurement of the work of each individual pupil. The average or median score of a class when considered alone may be very deceptive of the true conditions within that group as it gives no information with regard to the individuals measured. Every large school system represents relatively unselected groups of individuals ranging in intelligence from the subnormal to the near-genius type. To compare the performance of every pupil in a given test with grade norms may result in conclusions which are highly erroneous and unjust. The slow pupil is underrated and the superior pupil is overrated. Accurate and just measurements of the individual pupil necessitate measuring the performance of each pupil in terms of his own intelligence.

A beginning has been made in this direction by attempting to establish age, as well as grade norms. Buckingham and Monroe in the Illinois Examination have provided age norms. By dividing the achievement age of a pupil in reading by his mental age a measure of his reading ability is secured which is called the achievement quotient. Franzen in the Teachers' College Record of November, 1920, suggests a similar measure which he calls the accomplishment quotient.

A procedure for the measurement of the efficiency[1] of a pupil that allows the use of any standard test with the now available grade norms is herewith suggested. This method does away with the necessity of any computation as the quotients are quickly obtained by reference to tables.

---

[1] The writer is aware that this method does not give as accurate results as may be secured with the use of age norms and standard deviations. However, as only a relatively small number of the standard tests have been thus standardized the large majority of our tests are of little value for individual diagnostic purposes. When the method suggested by the writer is used one must keep in mind that the greater the deviation of the age of a given child from the average age of the grade the more unreliable the efficiency quotient becomes. Nevertheless in spite of the limitations of this method the writer is convinced of its practical value as a key to individual instruction.

What relationship exists between grade norms and the intelligence of the pupils tested? The standards for any reliable test are determined by being given to thousands of school children. After the scores for the various grades are compiled the median score for each grade is determined and used as a standard for the grade. This standard median represents the middle-most measure. The distribution of intelligence in this large unselected group forms a bell-shaped curve known as the probability curve or normal curve of distribution. Fifty percent of this group are of average intelligence and the median intelligence of the group represents the pupils with intelligence quotients of approximately 100.

Because of the high correlation between intelligence and scholarship the median point scores used as grade norms in reality represent the scores that have been made by pupils with intelligence quotients of approximately 100. The standard performance of a pupil with an intelligence quotient of 80 should be .8 of the grade norm while that of a pupil with an intelligence quotient of 120 should be 1.2 of the grade norm. Nevertheless, the important thing to be determined is not whether the performance of one pupil is equal to that of another in a given class or of some mythical average, but whether or not a pupil is doing all that he can be expected to do, because of his intelligence or inherent capacity for learning. The standard for each pupil should be an efficiency of 100 percent in terms of his own intelligence.

By dividing a pupil's point score in any test by the grade norm, which is based upon an I. Q. of 100, the performance of the pupil in comparison with the performance of the average pupil is determined. This measure is the pupil's achievement quotient. If the pupil's performance in any test is equal to the grade norm, his achievement quotient is 100. However, if his intelligence quotient is above 100, he is not doing the work that should be expected of him. In other words he is not 100 percent efficient.

The following data suggest the reliability of using grade standards as norms for pupils of intelligence quotients of 100. These data are taken from results secured by giving the Thorndike-McCall Reading Test to the fourth, fifth, and sixth grades in the West Allis Schools. Out of this group of 730 pupils, there were 120 with intelligence quotients from 96 to 104 inclusive. The medians of both the intelligence quotients and achievement quotients of these pupils were 100. In other words the median scores of the pupils in the fourth, fifth, and sixth grades with intelligence quotients of 100 were exactly the same as the grade norms.

By dividing the achievement quotient by the intelligence quotient a measure of the pupil's efficiency in terms of his own intelligence is determined. This measure is the efficiency quotient. The standard of efficiency for every pupil is 100. If the pupil's intelligence quotient is 110 and his achievement quotient 90, his efficiency quotient is 82.

Table I gives the median quotients of the several sections in spelling in the seventh and eighth grades in the Junior High School. The pupils are classified according to ability. The

TABLE I.—SPELLING. TEACHERS COLLEGE SEVEN S SCALES

Grade	Section	Point Score	Intelligence Quotient	Achievement Quotient	Efficiency Quotient
VII-B	1	48	78	78	100
	2	54	91	85	97
	3	64	104	102	98
	4	74	118	118	100
VII-A	1	54	81	85	105
	2	62	92	98	106
	3	66	110	105	95
VIII-B	1	52	92	72	78
	2	72	94	102	108
	3	70	100	98	98
	4	72	100	102	93
VIII-A	1	58	85	82	96
	2	58	100	82	82
	3	74	116	105	90

intelligence quotient forms the basis for this classification. The data are taken from a total of 330 pupils. Table II gives the median quotients in reading in Grades IV, V, and VI in West Allis.

TABLE II.—THORNDIKE-MCCALL READING TEST

Number of Pupils	Grade	Intelligence Quotient	Achievement Quotient	Efficiency Quotient
136	IV-B	91	98	108
109	A	89	97	109
141	V-B	90	94	104
102	A	97	99	102
149	VI-B	105	98	93
95	A	95	91	96

The situation as it may develop in the case of an entire class is indicated in Table III. This table presents the quotients for the five children in a fifth grade who had the highest intelligence quotients, and for the five children having the lowest intelligence quotients.

TABLE III.—FIFTH GRADE. THORNDIKE-MCCALL READING TEST

Pupils	Intelligence Quotient	Achievement Quotient	Efficiency Quotient
S. T.	140	127.5	93
J. B.	129	142.5	112
T. B.	126	132.5	108
K. B.	123	87.5	72
L. H.	121	132.5	112
L. C.	78	95	119
J. G.	77	87.5	120
O. N.	76	85	113
J. G.	72	80	106
R. I.	71	85	121

From these data it would appear that the slow pupil is more efficient than the superior pupil. Is this not a result of directing classroom instruction to the pupils of average intelligence and using every stimulus and incentive with the slow pupils to raise their standard of work? What has been done for the superior pupil? By putting forth a minimum amount of effort they have been doing as well as, or better than, the average pupil. They have been working under low pressure and have acquired habits of study and work that represent a low rate of efficiency when compared with their superior intelligence.

By classifying pupils according to intelligence and readjusting the subject matter and methods of instruction to the needs of each group, the work of each pupil can soon be placed on a high plane of efficiency. This may be illustrated by the following example. During the first semester of the school year 1920-1921 the sixth grade in one of the buildings was divided into two sections under separate teachers. The pupils were not classified according to ability. At the beginning of the second semester, however, the pupils of high average intelligence were grouped together under one teacher. Twenty-eight pupils were in this group and the intelligence quotient was 105. Table IV shows the quotients of this group for Monroe's Silent Reading Test. The first two tests were given before the pupils were classified. The last test was given after one semester's work in the classified section.

TABLE IV.—MONROE'S SILENT READING TEST. GRADE VI

Date of Testing	Achievement Quotient	Efficiency Quotient	Increase in Efficiency (percent)
September median..........	105	100	..
January median...........	120	114	14
May median..............	170	162	42

The reader's attention is called to the fact that during the four months that the pupils were unclassified, the increase in efficiency was but 14 percent but that during the four months from January to May—i. e., during the period when those of high average intelligence were grouped together—it rose from 114 to 162 with a percentage increase of 42.

Standard tests are of paramount importance to the classroom teacher as diagnostic instruments revealing the type of remedial instruction to be given. They may prove of inestimable value to the pupils themselves. If the results of the tests are placed before the pupils in a graphic form, it will be an effective means of motivation. An objective method of placing the performance of the pupils before them for comparison with standard goals will serve as a very wholesome incentive for more efficient work. Figure 1 is a reproduction of one of the graphs used for this purpose in the West Allis Schools. The achievement curves were based upon a list of twenty words taken from the Ayres Scale. Curve A shows the distribution of achievement quotients at the beginning of the semester. Curve B shows the distribution of quotients at the end of the semester. Space is also allowed for subsequent tests to be graphically recorded for comparison with Tests A and B. The median intelligence quotient of the class was 105.

Several general considerations may be mentioned in conclusion, considerations which affect the size of the efficiency quotient. Confining our attention to the conditions in virtue of which the efficiency quotient may rise above 100, we may observe that this result may be brought about in any one of four ways.

First, the original norms for the tests in question may be too low. They may be gathered from school systems where progressive measures and remedial activities have not been taken. To the extent that this is true, the published standards are too low. A pupil's achievement quotient will be represented as too high. Since his efficiency quotient is his achievement quotient divided by his intelligence quotient, there will thus be a fictitiously high representation of efficiency.

Second, due to non-intellectual qualities of a high order, the pupil in question may really be able to do better work than the average pupil of his intelligence. If such is the case, the efficiency quotient becomes a valuable measure of the extent to which these non-intellectual traits are operative.

Third, if dependence is placed upon grade norms for obtaining achievement quotients, a pupil may be represented as highly efficient because he is classified in a grade too low for his ability. Under these circumstances, he naturally scores high in the tests with reference to the standard of the grade to which he belongs. If, however, he were advanced to a higher grade, his performance would not yield as high an efficiency quotient.

Finally, an efficiency quotient above 100 may be the result of superior instruction or of superior administrative procedure. This is the inference to be drawn from Table IV.

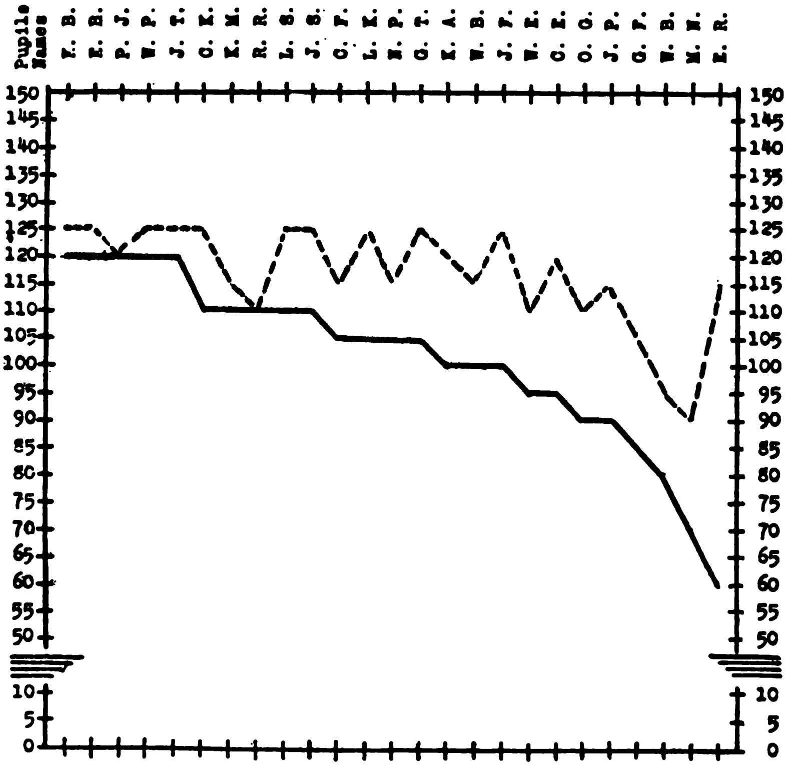
Test I. Curve A. Date, *January, 1921.* { **Median Achievement Quotient, *105***

Median Efficiency Quotient, *100*

Test II. Curve B. Date, *May, 1921.* { **Median Achievement Quotient, *120***

Median Efficiency Quotient, *115*

. . . . . .	VI-A	Spelling	. . . . . .
Building	**Grade**	**Subject**	**Teacher**



(Curve A, full line; Curve B, broken line.)

FIGURE 1.—DIAGNOSTIC ACHIEVEMENT RECORD

SUMMARY

1. Comparing the performance of a pupil with a standard grade norm does not give the proper measure of the pupil's work. Every pupil must be measured in terms of his own intelligence in order to determine his efficiency.
2. Standard grade norms represent the performance of pupils with intelligence quotients of 100.
3. Standard grade norms may be used as standards for the pupils of the given grade with intelligence quotients of 100.
4. A pupil's point score divided by the grade standard gives the achievement quotient which is a measure of the performance of the pupil in comparison with the performance of the average or normal pupil with an intelligence quotient of 100.
5. The achievement quotient divided by the intelligence quotient expresses the efficiency of the pupil in terms of his own intelligence. This measure is the pupil's efficiency quotient.
6. The quotients are found by reference to tables that have been prepared that embrace all standard tests.
7. A pupil's efficiency quotient may be above 100 due to:
    - a. Standard norm too low.
    - b. Pupil does better work than the average child of his intelligence.
    - c. Improper classification of pupil.
    - d. Superior instruction.
8. The average slow pupil is working at a higher rate of efficiency than the superior pupil.
9. The performance of the pupils displayed in graphic form will prove a wholesome incentive to the pupils for increased efficiency.

HOW MUCH MENTAL ABILITY DOES A TEACHER NEED?

W. B. Bliss
State Department of Education, Columbus, Ohio

The educational air is filled with tests and talk of tests but rarely is there occasion to test the testers. More and more teachers are coming into service almost entirely unmeasured except through the rather formal evaluation of the training-school instructors and the halting and discounted activities of the practice classes. Diplomas and certificates to teach have followed hastily in the wake of weeks checked off on the calendar of training, and we have faced the fact that we had with us a considerable number of young people who stepped forth to teach, casting the shadow of a well-trained teacher without the substance. This leads us to the conclusion, that training schools are not omnipotent, that they cannot make competent teachers out of less than competent human beings, nor open broad professional horizons to those who are left below the rim. It is commonly assumed that the training agency must build upon a certain amount of native ability, a composite of qualities of the heart and mind. It is the purpose of this paper to consider the latter somewhat, though there will be no pretense of a definite answer to the ambitious interrogatory of the title: How much mental ability does a teacher need? A few observations, however, will be offered on the subject.

The gifted individuals do not greatly exceed five percent of our total population (from five to seven percent, according to various estimates). In Ohio we have, therefore, in round numbers, perhaps 300,000 superior men and women. If we assume that the teaching profession draws 5 percent of these, we are undoubtedly liberal in the face of the competition of business and the well established professions. This would give to education not over 15,000 gifted people. Since the entire teaching force of the state is about 35,000 persons, this would leave about 20,000 others to be recruited from the great average class. In fact, if the investigation made at Chillicothe high school last year is an index to a general situation, we are far too liberal in the above estimate. The average intelligence quotient for the whole senior class was 102; for those who elected teaching it was 98; for those who elected law, medicine, and engineering it was 121. Four or five of the thirteen who expressed a desire to teach were below the average in ability and very unpromising. If the notion has been current in America that the schoolmaster or schoolmistress is an oracle and an outstanding paragon of knowledge and ability, that notion must be let slip in the face of the facts. Nor can we demand that the teacher shall be of superior mental ranking unless we at the same time have a magic carpet that will convey one-eighth of the state's gifted individuals straight to the fields of education. However, we do want to ask for more and more superior individuals and we must also strenuously insist that we shall never go below the group of average ability for any of our recruits.

Teaching requires too large a body of workers to secure only the mentally elect. No other profession in Ohio approaches the size of the teaching profession. The lawyers come closest with approximately 9000 enrolled. Of doctors there are about 6500, of dentists and pharmacists 4500. Here then is a vocation whose workers are four times as numerous as are those engaged in law, and five-and-a-half times as numerous as the doctors. It is beyond reason to hope that we shall soon cease to draw upon average people to fill the ranks of such a tremendous professional army, but we can expect a progressive development in the direction of a steadily increasing proportion of gifted recruits. That is our real hope. The teacher may not be a seer and a prophet nor satisfy Goldsmith's wonder "That one small head could carry all he knew," but she (the gender must be shifted) should be an educable human being of reasonable parts. It must be insisted that her mentality be greater than that of the 300,000 dullards which we possess as an offset to the like number of gifted.

The Department of Education made a survey last year of the students in the county-normal schools of the state through the medium of group intelligence tests. Evidence was wanted regarding the mental power of the entire group—the cerebral complexion, as it were. It was not found that the group of 743 students measured, ranged from average to superior, as our program calls for, but that they ran the gamut from inferior to gifted. Quite naturally the query was: Is there not a line below which not enough mental ability exists to make the individual at all promising in spite of other desirable qualities? The answer seemed to be in the affirmative. Those who analyze the necessary qualities that go to make up the successful teacher are quite sure that mental ability is only one factor, but they do not deny that when it is present in only very limited extent it negatives and renders abortive any other ordinarily essential qualities. The Department's premise, therefore, has been that success as a teacher may not correspond perfectly to mental ability but that there is a vanishing point where absence of mental alertness will be followed by lack of teaching success.

The Terman Group Test of Mental Ability was administered in 42 county-normal schools last year. The distribution of scores approximated the normal curve, revealing, as pointed out above, that persons of all grades of ability were entering the teaching profession; and there seemed no ground for assuming that any strongly modifying selective process had been at work.

The median score was 141.86. The quartiles were 117 and 162, meaning that half of the students made scores between these two figures. In any large unselected group we should expect sixty percent to be of normal, or average intelligence, so applying this principle to the data secured by these tests, we can say that those who made scores between 112 and 168 are of average mental capacity. It naturally takes some individuals somewhat below the exact median and some above it to produce an average and those at or near 112 have not the power of those at or near 168, but within those confines are found the varying grades of power which characterize the broad plateau of mediocrity.

During the first part of this current school year (1921-22) the same test was again administered in 49 county-normal schools to 1305 pupils. The results show an interesting conformity with last year's data. The median was 141.21 as against 141.86 cited for last year. The other points on the distribution compare as follows:

	Last Year	This Year
7-Percentile....	92	92
20-Percentile....	112	112
25-Percentile....	117	119
50-Percentile....	142	141
75-Percentile....	162	162
80-Percentile....	168	167

The equivalence of the two distributions was very significant and strengthens faith in the whole program. The 7-percentile, which was 92 both years, has been taken as the vanishing point, referred to before. Students scoring at or below this point have been singled out for elimination from the training schools unless their score on another reliable group test contradicted the Terman score. The seven percent mark was chosen because that point is commonly accepted as cutting off the distinctly questionable portion of any distribution not too highly selected.

The Illinois General Intelligence Test was given to 900 of the same students who took the Terman Test. The data secured with the Illinois scale were as follows:

7-Percentile.....	87
20-Percentile.....	101
25-Percentile.....	104
50-Percentile.....	119
75-Percentile.....	133
80-Percentile.....	137

In some cases the National Intelligence Tests were used, in some the Otis and in others the Army Alpha, the aim always being to have more than one measure of a student.

Out of the 1305 students 55, or slightly over 4 percent, were advised to take up other lines of work rather than continue preparation for teaching. Of the 55 deemed undesirable prospects for teaching 51 took the Terman Test. The median score was 77; the average 75. Thirty-seven took the Illinois with a median score of 79 and an average of 77. The Department has had reports thus far to show that 44 students have withdrawn from county normals in line with our suggestion; 42 of these had taken the Terman Test with a median for the group of 78. Thirteen students were tested with the Stanford-Binet individual examination; the median I. Q. was 80, the average 78.

A comparison was made last year between the students' test scores and their high-school marks. The grades of each student were averaged for the full four years of the high-school course and a correlation obtained between these figures and the test results. The coefficient of correlation was 0.43. This may be considered as a marked correlation but does not establish a very close correspondence between ability as revealed by the test and quality of high-school work.

The foregoing facts are of statistical interest only. They do not throw light on our attempted discussion of how much ability a teacher needs. In the interest of securing certain data on this, a report was asked for from twenty-five county superintendents covering the teaching success of certain designated individuals who graduated from county-normal schools last year. Reports were received from 21 superintendents covering 77 individuals. In selecting those for whom reports as to success were to be given, the method in most cases was to designate the two who had made the highest scores in the normal school of the particular county last year, two or three of those who scored near the median of 142, and the two lowest. On a five-point rating scale these would be rated as to mental ability as 1, 3, and 5 respectively. Superintendents were asked to rate them as to teaching success with the figures 1, 2, 3, 4, 5; it was explained that these figures might be understood as signifying superior, excellent, good, fair, and poor respectively. It happened that the scores of two who were selected on the ground of being highest in their schools were not high enough to deserve the mental rating 1, so 2 was assigned in these cases; one was somewhat above the group who were rated 5, so 4 was assigned as a mental rating in this case.

The rating of the superintendents as to success agrees exactly with the rating as to mental ability in 20 cases out of the 77. This is a percent of 26, and may be taken for practical purposes as a measure of the correlation that exists.

A more compact showing as to the relation between the mental ratings and success in teaching as judged by the superintendents is disclosed by Table I.

The distribution of the 77 selected cases according to mental ability was as follows:

MENTAL ABILITY.......	1	2	3	4	5
NUMBER OF CASES.....	21	2	32	1	21

The distribution of these same cases according to teaching success was as follows:

TEACHING SUCCESS.....	1	2	3	4	5
NUMBER OF CASES.....	4	24	33	13	3

These figures are on the whole encouraging. They show that the chances are twice as good to have superior or excellent teaching if individuals of superior ability are employed, as when individuals of average ability are employed, and almost four-and-a-half times as good as when individuals of low rank mentally are hired. They show further that there is only one chance in ten of an individual of superior ability doing work that is as low as fair or poor, but there is one chance in four that average individuals will do grades of work this low. Perhaps the most surprising thing about these reports is that 57 percent of those who were at the bottom of the scale on the tests are doing good work and 14 percent are doing excellent work. None of the low group were rated as doing superior work and, conversely, none at the top of the scale mentally were rated as doing poor teaching.

TABLE I.—RELATION BETWEEN MENTAL RATINGS AND SUCCESS RATINGS

MENTAL RATING	RATING FOR SUCCESS IN TEACHING (PERCENTS)		
	"Excellent" or "Superior"	"Good" or higher	"Fair" or "Poor"
5	14[a]	71	29
3	31	75	25
1	61	90	10[b]

[a] None "Superior."
[b] None "Poor."

One point needs to be brought forward here and that is that the results on one mental test are not always conclusive. This year 91 individuals made scores at or below 92 on the Terman Test, but only 51 of these were singled out for elimination because it was found that their showing on the Illinois or some other second test did not corroborate closely enough the verdict of the first test alone. Hence, only 56 percent of those below the 7-percentile mark on the Terman Test were conclusive cases. This same percent may be assumed to have been valid last year, and that would mean that we are treating in this part of our study some students as 5 in mental ability who might have been found to be at least one notch higher. Applying this percent to the 21 of mental ability 5 treated in this part of the investigation, we should have about 12 students who could be considered incontestably 5. The chances are that with such a correction the success showing of the group would have been more in accord with reasonable expectations. It will be noted that while the Terman median for the group rejected this year was 77, it was 83 for the low group of last year—the group whose success we are studying; the averages were 75 and 81 respectively. Therefore our group for the purposes of this success study was somewhat above the level of those marked this year for elimination.

We might as well understand, however, that we cannot eliminate by these suggested revisions the disquieting fact that a large proportion of those lowest on the test are considered to be doing good work or better. It leads to the belief that either the tests or the ratings or both are wrong. Or if both the tests and the success estimates are dependable, we are led to the conclusion that a teacher does not need to be up to the average in mental ability. That is to say, our hypothesis is wrong; and there is no point below which qualities of address, of temperament, and of volition cannot compensate for lack of intellectual power. Some analysis is in order.

I am inclined to place considerable reliance upon the validity of the Terman Group Test. I recognize the limitations of group tests as absolute measures, but they do succeed in establishing relative degrees of ability to a highly satisfactory extent. I have already allowed for the fact that because this was the only test given last year some students may have been placed lower than a second test would have indicated, and that this may have been true in as much as 44 percent of the cases.

As to the superintendents' ratings I think that even they themselves would agree that there is greater possibility of error here.

The reasons are: (1) that there were 21 different measuring sticks used in these success ratings, all largely subjective, whereas there was only one objective measuring scale used in deriving the mental ratings; (2) that the item which I asked to be rated—"success as a teacher"—was so general and undefined as to invite a lump estimate, with some factors inevitably getting different weights with different people; (3) that I allowed so little time for reply that there was no opportunity to base the estimate upon a very extensive observation. For instance, a girl who was tested last year with the Stanford-Binet individual examination showed a mental age of 11 years and 5 months; the superintendent says her success is good. A girl who scored 67 on the Terman Test, and who was rated by the normal director last year as being 4 in general ability and 3½ in classroom work, is rated by the superintendent as 2 in teaching success. Another scored 91, was rated by the normal director 4½ in general ability, 4 in classroom work but 2 in success by the superintendent. These success ratings seem open to further investigation. There is a common human tendency, manifested by both directors and superintendents, to rate too high. For example, in rating the students in the normal schools last year the directors placed almost twice as many students too high as too low, if we use the test scores as a standard of reference. As far as the superintendents' high ratings are concerned, it may doubtless be said—and the statement really constitutes a tribute to the county normal schools—that the superintendents quite uniformly find this training making itself effectively felt and that even the weaker individuals stand out reasonably well when occasional sections of their work are seen in the early part of the year.

Since there is the possibility of errors both in the testing and in the superintendents' judgments and since the group studied is very small, we cannot hope to draw any very definite conclusions. An adequate comparison could only come through having one individual rate the whole group with very definite standards in mind as to what constitutes teaching success.

It is a rather common belief, to which I have subscribed, that the primary teacher does not need to have as much mental ability as teachers of other grades. Accordingly, normal directors and superintendents have at times suggested keeping in school those students who make low test scores and placing them in primary work. The only way to answer that question would be to measure a large number of primary teachers and see if the majority of those considered satisfactory were below average mentally. During Camp Sherman days Prof. C. A. Murchison administered the army test to the teachers of a certain city. It was found that the level of ability of the first-grades teachers was lower than that of any other group of grade teachers, but it was not found that any of them was below average. Seventy-six teachers were examined: 69.6 percent (an unusually high proportion) were rated very superior; 21 percent were superior, 7.8 percent were high average, and 1.3 percent were average. So while it is likely that the keenest and quickest mentalities are not often accompanied by those qualities of patience and sympathetic appreciation of halting efforts that function so well in primary work, it is unfair to assume that we must go below average ability to get good primary teachers, or even that adequate primary teaching can be achieved by individuals as low on the scale of relative ability as are those we have sought to eliminate this year.

Can we afford to develop our educational workers out of anything but good materials or can we develop the progressive, growing type otherwise? In an age which demands mental keenness and alertness for success in any important calling we cannot afford to demean the teaching profession by demanding anything less. It is apparent from this study that our best prospects for success are to be found in individuals of above-average mental ability. Therefore, while keeping in mind that we cannot find enough superior people to fill the profession, we can constantly strive to get more and more of this better type. We need not gainsay the merit of the teacher of average ability who possesses to a marked degree many wonderful qualities that enter into the successful teacher's make-up. But we are sure to be wrong if we minimize mental ability, if we expect other qualities to offset its lack, and if we do not stand for the application of some selective process by teacher-training institutions. This limited investigation shows that you can expect teaching success to keep reasonable pace with mental ability.

COMPARISON OF THE BINET-SIMON AND OTIS TESTS

S. C. GARRISON AND JAMES S. TIPPETT
Peabody College for Teachers

The purpose of this article is to make a comparative study of results secured by giving the Stanford Revision of the Binet-Simon Test and the Otis Advanced Group Intelligence Test to 158 pupils of the Peabody Demonstration School.

The Peabody Demonstration School is conducted in connection with George Peabody College for Teachers. Several contributing factors give this school a somewhat selected enrollment. A rather large tuition fee is charged; therefore most of the children come from the homes of the business and professional classes. Pupils who are unable to do the work of the grades successfully, either because of poor ability or because they cannot be induced to make further earnest effort, are asked to withdraw from the school. A further selective measure is the fact that during the last two years pupils who have been admitted to the school have been required to take a psychological examination. If their ability has been found to be poor, they have not been allowed to enter.

During the months of February and March, 1921, all the children (158) of the fourth, fifth, sixth, seventh, and eighth grades of the Peabody Demonstration School were given the Stanford Revision of the Binet-Simon Test. One month later the same children were given Form A of the Otis Advanced Group Intelligence Test. The tests were given by the professor of educational psychology in George Peabody College for Teachers or, under his direction, by graduate students in the department of educational psychology. In administering the tests, the directions were followed minutely and the scores were carefully checked. Many of the children have taken the Binet-Simon Test two or more times and their successive records have been found to agree in a surprising manner.

Table I gives the distribution of the pupils according to actual ages, Binet-Simon Ages, and Otis Mental Ages. The ages have been tabulated so that 9 years means from 9 years to 9 years and 6 months. The median actual age for the 158 children was 11 years and 10 months. The median Binet-Simon Mental Age was 12 years and 4 months. The median Otis Mental Age was 14 years and 1 month. Thus the Binet-Simon Test gives a median I. Q. of 104; the Otis Test gives a median I. Q. of 118.

TABLE I.—DISTRIBUTION OF PUPILS ACCORDING TO ACTUAL AGE, BINET-SIMON MENTAL AGE, AND OTIS MENTAL AGE

YEARS	FREQUENCY ACTUAL AGES	FREQUENCY BINET-SIMON MENTAL AGES	FREQUENCY OTIS MENTAL AGES
8	0	0	4
8	0	0	0
8½	0	5	6
9	10	1	2
9½	6	1	4
10	16	9	10
10½	16	13	5
11	17	25	3
11½	15	11	8
12	16	18	19
12½	18	6	8
13	18	10	9
13½	15	7	7
14	6	8	16
14½	5	5	3
15	0	11	11
15½	0	6	7
16	0	7	1
16½	0	7	1
17	0	3	4
17½	0	2	2
18	0	3	38
MEDIAN	11 yrs. 10 mo.	12 yrs. 4 mo.	14 yrs. 1 mo.

The median Actual Age, the Binet-Simon Mental Age, the Otis Mental Age, and the difference in months between the Binet-Simon and Otis results for the different age groups are shown in Table II. An examination of the table shows that for every age group the Binet-Simon median is smaller than the Otis median. The difference between the medians for the 9- and 10-year groups is rather small; the Otis medians being two and six months respectively greater than the Binet-Simon medians. When we examine the 11-year group, however, and the other groups in the table, we find rather large differences: 11 months for the 11-year group, 20 months for the 12-year group, 30 months for the 13-year group, and 28 months for the 14-year group.

TABLE II.—MEDIAN AGES FOR EACH OF THE YEAR GROUPS

YEAR GROUP	CASES	MEDIAN ACTUAL AGE	MEDIAN BINET-SIMON MENTAL AGE	MEDIAN OTIS MENTAL AGE	DIFFERENCE BETWEEN B-S AND OTIS (Months)
9	16	9-6	10-9	10-11	2
10	32	10-6	11-4	11-10	6
11	32	11-5	11-9	12- 8	11
12	34	12-5	13-3	14-11	20
13	33	13-5	15-2	17- 8	30
14	11	14-4	15-4	17- 8	28

Table III shows for each grade the highest, median, and lowest mental ages for both the Stanford Revision and the Otis Test. The table shows considerable variability with respect to these age limits. The highest and lowest Binet-Simon mental ages for each grade are with one exception (Grade V) closer together than the highest and lowest Otis mental ages.

TABLE III.—THE MEDIAN, HIGHEST, AND LOWEST OF THE BINET-SIMON AND THE OTIS MENTAL AGES BY GRADES

GRADE	CASES	BINET-SIMON MENTAL AGE			OTIS MENTAL AGE		
		Median	Highest	Lowest	Median	Highest	Lowest
IV	35	10- 9	12-9	8- 8	10- 5	16-2	8-
V	33	11- 9	18+	9- 5	13- 4	18+	9-9
VI	34	12- 4	17-9	10- 4	14- 3	18+	10-3
VII	27	14-10	18+	11- 7	15-10	18+	12-7
VIII	29	15-11	18+	10-11	18+	18+	9-9

The tables presented thus far show considerable differences between the results secured by the two tests. This is further shown by Tables IV and V.

Table IV shows, for the actual age groups, the frequency of difference in months between the mental ages given by the Binet-

Simon Test and the Otis Test. The differences are grouped in steps of six months. For example, five pupils of the 9-year actual age group when examined by the two tests were given mental ages varying less than six months, while two were given mental ages varying more than six months but less than twelve months.

TABLE IV.—DISTRIBUTION OF DIFFERENCES BETWEEN BINET-SIMON AND OTIS MENTAL AGES FOR EACH AGE GROUP

DIFFERENCE IN MONTHS	ACTUAL AGE GROUPS						TOTALS
	9 Years	10 Years	11 Years	12 Years	13 Years	14 Years	
0– 6	5	6	7	7	5	1	31
–12	2	10	7	5	6	1	31
–18	3	1	5	4	7	3	23
–24	1	6	2	1	3	4	17
–30	1	3	5	3	2	1	15
–36	1	1	4	5	3	1	15
–42	2	2	1	2	2	0	9
–48	0	0	0	1	1	0	2
–54	0	2	0	1	3	0	6
–60	0	1	0	2	1	0	4
–66	0	0	0	1	0	0	1
–72	0	0	1	1	0	0	2
–78	0	0	0	1	0	0	1
–84	1	0	0	0	0	0	1
CASES..........	16	32	32	34	33	11	158

The average difference in months between the mental ages given by the two tests and also the average plus difference in favor of the Otis test for the actual age groups is shown in Table V. There is an average plus difference in favor of the Otis test of more than a year for each of the age groups while

TABLE V.—AVERAGE DIFFERENCES BETWEEN BINET-SIMON AND OTIS MENTAL AGES

AGE GROUP	AVERAGE DIFFERENCE IN MONTHS	AVERAGE PLUS DIFFERENCE IN FAVOR OF OTIS
9	19.4	14.3
10	19.4	13.6
11	18.7	13
12	26	24.7
13	22.7	20.7
14	20	15
ALL	21.5	17.5

there is an average difference between the two tests of more than one and one-half years.

The average difference in months for any age in Table V was found by getting the sum of all differences between the Binet-Simon result and the Otis result for the cases in the age group, and then dividing the sum by the number of cases. The average plus difference in favor of the Otis Test was found by dividing the difference between the sum of the plus differences for any actual age group and the sum of the minus differences for that age group by the number of cases in that age group.

It so happened that a survey of the school by means of educational tests was made a few weeks before the mental tests were given. The following tests were used: the Woody-McCall Fundamentals in Arithmetic Test, the Monroe Reasoning Test in Arithmetic, the Ayres Spelling Scale, the Thorndike Handwriting Scale, the Hillegas Composition Scale, and the Monroe Silent Reading Test. The pupils in each grade were ranked from best to poorest for each test. A composite ranking was made from these ranks by adding all the ranks which each pupil attained and re-ranking these sums from smallest to greatest.

The teachers of each grade were also asked to rank their pupils from best to poorest. They were directed to base their ranking not on school work alone but on general ability as displayed by the pupils on the playground and in the home, as well as in school.

After these rankings were made, coefficients of correlation (rank order method) were worked out for each grade as shown in Table VI. In four of the five grades the Otis Test correlates higher with the teachers' ranking than does the Binet-Simon Test. The Otis Test correlates much higher with the ranking given by the educational tests than does the Binet-Simon Test. The teachers' ranking and educational test ranking give a much higher coefficient of correlation than do either of the mental tests with one of these rankings. When the Binet-Simon and Otis results are correlated we secure rather lower correlations than would be anticipated. However, when the entire 158 cases are correlated

TABLE VI.—COEFFICIENTS OF CORRELATION BETWEEN THE BINET-SIMON TEST, THE OTIS ADVANCED TEST, THE TEACHERS' RANKING, AND THE EDUCATIONAL TESTS RANKING

GRADE	CASES	RANKINGS BY	BINET-SIMON STANFORD REVISION	OTIS ADVANCED GROUP	TEACHERS' RANKING	EDUCATIONAL TESTS RANKING
IV	35	Binet-Simon, Stanford Revision		54	67	62
		Otis Advanced Group	54		74	78
		Teachers' Ranking	67	74		62
		Educational Tests Ranking	62	78	62	
V	33	Binet-Simon, Stanford Revision		57	45	43
		Otis Advanced Group	57		61	73
		Teachers' Ranking	45	61		84
		Educational Tests Ranking	43	73	84	
VI	34	Binet-Simon, Stanford Revision		54	54	33
		Otis Advanced Group	54		83	62
		Teachers' Ranking	54	83		71
		Educational Tests Ranking	33	62	71	
VII	27	Binet-Simon, Stanford Revision		23	38	27
		Otis Advanced Group	23		25	59
		Teachers' Ranking	38	25		74
		Educational Tests Ranking	27	59	74	
VIII	29	Binet-Simon, Stanford Revision		53	54	32
		Otis Advanced Group	53		67	54
		Teachers' Ranking	54	67		69
		Educational Tests Ranking	32	54	69	
AVERAGES		Binet-Simon, Stanford Revision		48	52	39
		Otis Advanced Group	48		62	65
		Teachers' Ranking	45	61		84
		Educational Tests Ranking	43	73	84	

without regard to grade, a coefficient of correlation of 75 between the Otis Test and the Binet-Simon Test is found.

Although the number of cases included in this study is too small to permit any very general conclusions, several things seem to be clearly indicated:

(1) Mental Ages secured by means of the Binet-Simon Test and the Otis Advanced Group Test cannot be used interchangeably.

(2) The Otis Test tends to scatter the scores more than the Binet-Simon Test.

(3) The Otis Test tends to give a higher mental age than the Binet-Simon Test.

(4) The Otis Test and the Binet-Simon Test give a fairly high coefficient. (0.75, n = 158, Grades IV-VIII.)

(5) The Otis Test correlates higher with educational test rankings and teachers' rankings than the Binet-Simon Test.

A DATA SHEET FOR THE PEARSON CORRELATION COEFFICIENT

L. L. THURSTONE
Carnegie Institute of Technology

I shall describe a data sheet which facilitates the calculation of the Pearson correlation coefficient. If one does not calculate the correlation coefficient often one may find it necessary to return to the textbooks in order to relearn the method of computation. That should not be necessary if a data sheet is available with direct instructions for the computation separated from the discussion of the logic of the correlation table. I find that even those who calculate these coefficients frequently spread their calculations over many slips of paper and compute the various parts in more or less random order. It is almost impossible to check calculations that are not systematically arranged. When I encounter calculations that have been made without a plan and without orderly arrangement I find it difficult to consult them. It is not uncommon to find correlation calculations which are not even labeled as to what the variables are. Such calculations are just as good as if they had been lost or misfiled.

The correlation data sheet which I have been using is of double letter size so that it can be folded once and filed with letter size paper and reports. I have arranged the sheet so that it should not be necessary to do any scribbling on loose slips of paper. A small space is provided for the scribbling that may be necessary. All items in the calculation, such as frequencies at the intersections of the arrays, sums, squares, and other numbers, belong in definite squares on the data sheet. If one has occasion to return to the calculations, one can readily find what one is looking for. The preparation of this sheet is not a contribution to statistical theory. It is merely a labor saving office device.

INSTRUCTIONS FOR THE USE OF THE DATA SHEET

In order to make these instructions brief and easily read I have indicated several of the columns by the numbers at the top of the sheet. These numbers do not occur on the regular data sheet. They are inserted here for the purpose of these instructions only. The horizontal rows are labeled by letters for the same purpose. The rows and columns are known collectively as arrays. When I refer in these instructions to a certain square on the data sheet I shall refer to it by the number of the column and the letter of the row which intersect at the square to which I wish to call your attention. Thus C-3 refers to a square on the sample data sheet which contains the figure 7. The sample set of data is drawn from West's Textbook in Statistics.[1]

(1) Select any suitable class interval for the X-variable. On the data sheet these intervals are unity.

(2) Record these class intervals in row A as shown (61, 62, etc.) Always arrange the small numbers to the left, and the large numbers to the right.

(3) Row B should contain the upper class limits for the class intervals, which in the case of the example would be 61.9, 62.9, etc. Row A contains, then, the lower class limit, and row B the upper class limit.

(4) Select a suitable class interval for the Y-variable. In the example this is five. Record the lower and upper class limits for these class intervals in columns 1 and 2 respectively. By tabulating both the upper and the lower class limits for each class interval one is sure that the actual range for each class interval is indicated on the data sheet. This serves to prevent confusion as to the square in which any particular frequency is to be recorded. Arrange the small numbers at the bottom and the large numbers at the top as shown.

(5) Record the frequencies in the squares of the correlation table. In the example these range between rows D and K and between columns 3 and 7.

(6) Record in row L the sums of the frequencies in the vertical columns of the correlation table.

(7) Record in column 8 the sums of the frequencies in the rows of the correlation table.

---

[1] West, Carl J. *Introduction to mathematical statistics*. Columbus, Ohio: R. G. Adams and Company, 1918. Page 67.

(8) Add the numbers in the row L and record at N (N =) in the table of sums in the upper right hand corner of the data sheet. In the example this sum is 750. This is the total number of cases.

(9) Add the numbers in columns 8 and see if the sum agrees with the N just recorded in the table of sums. It should be the same number.

(10) Select any class interval as an arbitrary origin for the X-variable. In the example this is chosen at column 5. This arbitrary origin may be taken at any class interval but it reduces the arithmetical labor to choose it as close to the median as possible. This should be done by inspection because it does not pay to calculate the median specially for this purpose. Draw a blue pencil line or a wavy pencil line through the column 5 so as to mark off clearly at all parts of the correlation table the position of the arbitrary origin.

(11) Do likewise for the Y-variable. The arbitrary origin for the Y-variable has been chosen at row G. This row is marked similarly to the corresponding class interval in X (column 5). The correlation table is now divided by the blue pencil lines into four quadrants. It should be remembered that these quadrants are marked off from the two assumed means and are not to be confused with the quadrants which would be marked off from the true means. No confusion will result unless one should attempt to calculate the coefficient by the method of unlike signs by means of the assumed quadrants.

(12) Record a zero at M-5. Record the successive class intervals on each side of this zero as shown in the rest of row M. Do likewise in row C.

(13) Record in row N the products fx. These are merely the products of the two rows immediately above N.

(14) Record in row P the products fx^2. These are the products of the two rows immediately above P. This is easily verified by inspecting the sample data sheet.

(15) Record a zero at G-9. Record above and below this the ascending and descending orders of class intervals as shown in the rest of column 9.

(16) Record the products fy in column 10. These are the products of the two columns immediately to the left of column 10.

(17) Record the products fy^2 in column 11. The details of these calculations should be evident from the data sheet.

(18) In row D you will find the numbers 3, 2, 2, 1 to the right of the assumed origin. Immediately above these numbers in row C you will find the numbers 1, 2, 3, 4. Each of the numbers in row D is to be multiplied by the number immediately above it in row C. These numbers are to be summed up thus:

$$(3 \times 1) + (2 \times 2) + (2 \times 3) + (1 \times 4) = 17.$$

The number 17 is recorded at D-13. The sums that are obtained on the right hand side of the assumed origin (column 5) are recorded in column 13. The sums similarly obtained on the left hand side of the origin are recorded in column 12. This verbal statement will be clearer by noting visually on the data sheet the manner of the calculation.

This operation is repeated for each row with a separate sum for the right hand side and for the left hand side. For example: In row F we have the following sums.

Left side: $(3\times1)+(2\times2)+(2\times3)=13$
Right side: $(10\times1)+(12\times2)+(11\times3)+(1\times5)+(1\times6)=78$
These sums, 13 and 78, are recorded in the squares F-12 and F-13 respectively.

This operation can be performed with less effort on a comptometer or any other key driven calculating machine. It is done as follows:

Consider the left side of row F. Set the finger on the 1-key at the lower right end of the machine. This key represents the deviation 1 of the C-row. Punch the key three times to represent the frequency of three. Move the finger to key 2 immediately above your present location on the machine. Punch it twice to represent the frequency of two. Move the finger up one step again and punch the new key twice to represent the third frequency of two. The sum of 13 is now on the machine without any further mental effort. This work can be done without looking at the machine so that one may keep one's eyes on the correlation table.

Now consider the right side of row F. Place the finger on the 1-key at the lower right end of the machine. This key may be punched ten times or the key to its left may be punched once to represent the frequency of ten. Move the finger up one step on the machine. Punch this key twelve times or punch it twice and the key to its left once to represent the frequency of twelve. Move the finger up one step on the machine and punch the key eleven times or punch it once and the key to its left once. Move the finger up two steps on the machine and punch once; move the finger up once more and punch once. The sum of 78 is now in the machine and may be recorded at F-13. One will readily learn to move the finger one step at a time on the machine without looking at it even though one may not know more about the calculating machine. In this way the task is rendered extremely easy even for a novice.

(19) The difference between the two numbers 13 and 78 which were recorded at F-12 and at F-13 is 65. This number 65, is recorded at F-15. The same operation is carried out for all the figures in columns 12 and 13. Their differences are recorded in columns 14 and 15. If the number in the 12-column is larger than the adjacent number in the 13-column one records the difference in the 14-column. Similarly, if the number in the 13-column is larger than the adjacent number in the 12-column one records the difference in column 15. One should visualize this right and left relation in order to remember it.

(20) For the portion of the table above the assumed origin of the Y-variable (row G) record the products of the corresponding numbers in columns 9 and 14 in column 16. Record the products of the corresponding numbers in columns 9 and 15 in column 17. These instructions apply above the assumed origin. Below the assumed origin the relation is reversed so that the products 9 and 14 are recorded in column 17 while the products 9 and 15 are recorded in 16. The following summary will make matters clearer.

Columns
Above the assumed origin: (9)×(14)=(16)
(9)×(15)=(17)
Below the assumed origin: (9)×(15)=(16)
(9)×(14)=(17)

Another way to recall this is to note that columns 14 and 15 preserve the same spatial arrangement above the assumed origin but that below the origin this lateral arrangement is reversed.

Still another and perhaps the best way to recall this is to note that the y-deviations in column 9 are positive above the assumed origin and negative below the assumed origin. This comes about from the fact that we plot the Y-variable in column 1 with the high numbers at the top of the column and the small numbers at the bottom. If now the positive y-deviations in column 9 are multiplied by positive Σx in column 15 we obtain a positive Σxy and record the product in column 17. If we multiply the negative y-deviations of column 9 by the positive Σx in column 15 we obtain a negative Σxy and record the product in column 16. If we multiply the positive y-deviations of column 9 by the negative Σx in column 14 we obtain a negative Σxy which we record in column 16. If we multiply the negative y-deviations of column 9 by the negative Σx in column 14, the product Σxy is positive and is recorded in column 17.

(21) We are now ready to extract the sums for the final computation of the coefficient. Add the numbers to the left of the assumed origin in row N. Record the sum at $\Sigma x_{neg} =$ in the table of sums in the upper right hand corner. It is 744.

(22) Add the numbers to the right of the assumed origin in row N. Record the sum at $\Sigma x_{pos} =$ in the table. It is 669.

(23) Determine the difference between Σx_{pos} and Σx_{neg} and record this difference at $\Sigma x =$ in the table of sums. It is —75. The relation is $\Sigma x = \Sigma x_{pos} - \Sigma x_{neg}$. Record the proper sign for the Σx.

(24) Add the numbers in the row P on both sides of the assumed origin. Record this at $\Sigma x^2 =$ in the table (4201).

(25) Add the numbers above the assumed origin in column 10 and record the sum at $\Sigma y_{pos} =$ in the table of sums (880).

(26) Add the numbers below the assumed origin in column 10 and record the sum at $\Sigma y_{neg} =$ in the table of sums (933).

(27) Determine the difference between Σy_{pos} and Σy_{neg} and record the difference at $\Sigma y =$ in the table of sums (—53). The relation is $\Sigma y = \Sigma y_{pos} - \Sigma y_{neg}$. Record the proper sign for the Σy.

(28) Add all the numbers in column 11. Record this sum at $\Sigma y^2 =$ in the table of sums (7377).

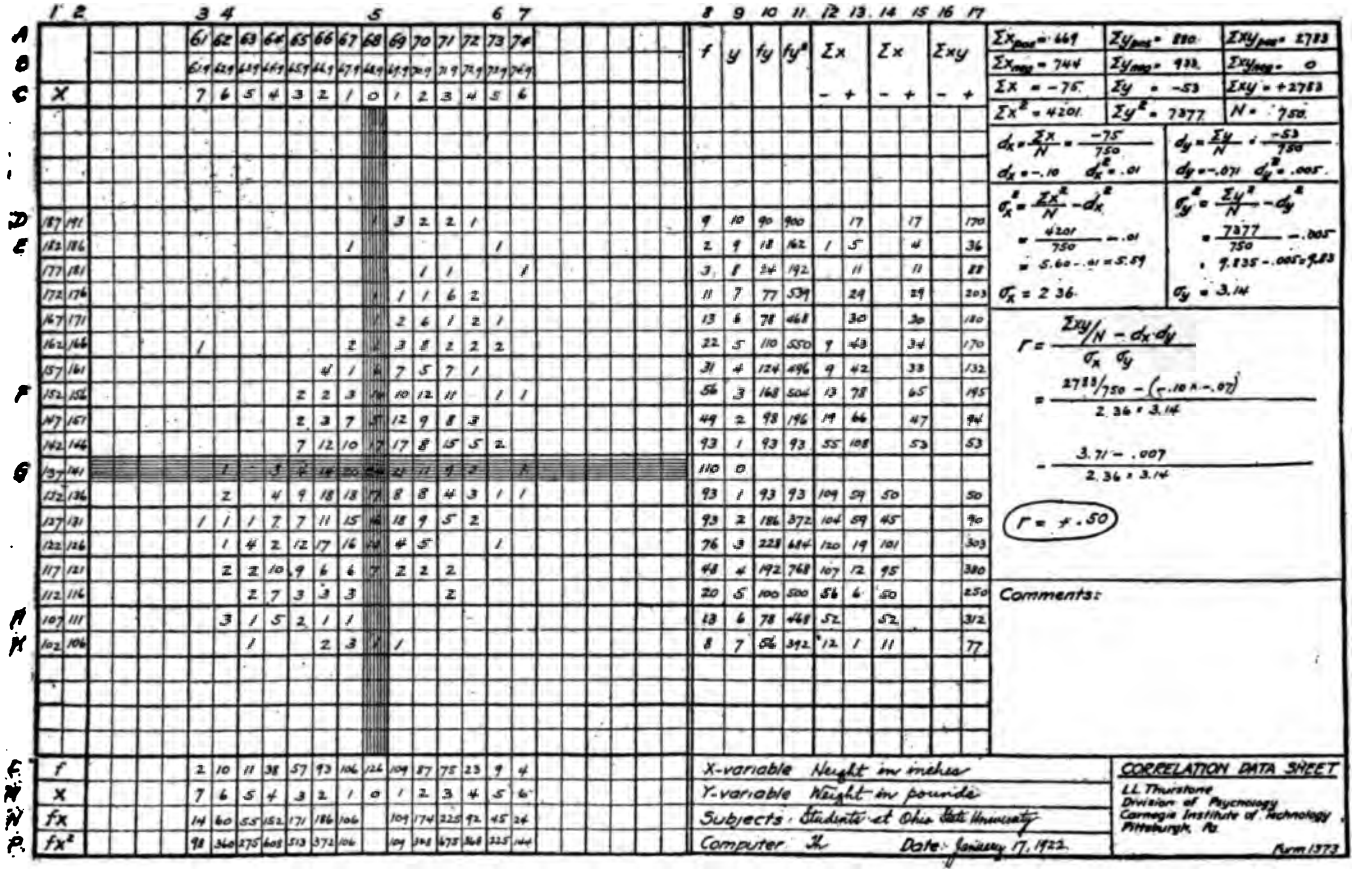
(29) Add all the numbers in column 16 and record at $\Sigma xy_{neg} =$ in the table of sums (zero).

(30) Add all the numbers in column 17 and record at $\Sigma xy_{pos} =$ in the table of sums (2783).

(31) Determine the difference between the Σxy_{pos} and the Σxy_{neg} and record the difference at $\Sigma xy =$ in the table of sums (+2783). The relation is $\Sigma xy = \Sigma xy_{pos} - \Sigma xy_{neg}$. Record the proper sign for the Σxy.

The numbers that have been recorded in the table of sums may be used for substitution directly into the formulae that are printed on the data sheet. From this point on one proceeds as shown by the formulae, turning the crank algebraically as it were, until the coefficient is obtained. In the example the correlation coefficient is +.50.

The work of tabulating the correlation table and computing the coefficient for a problem like the one in the accompanying example requires about forty to sixty minutes after one becomes accustomed to the sheet. If one is working with correlation coefficients continually, the calculation can be made in less time. Verbal description of spatial relations is necessarily clumsy. To read these instructions without consulting the data sheet might give one the impression that these calculations are awkward and even intricate; but after one has translated the necessarily clumsy verbal description into visual or graphic terms, the work can be done by a clerk who knows nothing about correlation statistics.



	1	2	3	4					5					6	7	8	9	10	11	12	13	14	15	16	17
A			61	62	63	64	65	66	67	68	69	70	71	72	73	74	f	y	fy	fy²	Σx		Σx		Σxy	
B			61.9	62.9	63.9	64.9	65.9	66.9	67.9	68.9	69.9	70.9	71.9	72.9	73.9	74.9										
C	X		7	6	5	4	3	2	1	0	1	2	3	4	5	6					−	+	−	+	−	+
D	187	191								1	3	2	2	1			9	10	90	900		17		17		170
E	182	186							1						1		2	9	18	162	1	5		4		36
	177	181										1	1			1	3	8	24	192		11		11		88
	172	176								1	1	1	6	2			11	7	77	539		29		29		203
	167	171								1	2	6	1	2	1		13	6	78	468		30		30		180
	162	166	1						2	2	3	8	2	2	2		22	5	110	550	9	43		34		170
	157	161						4	1	6	7	5	7	1			31	4	124	496	9	42		33		132
F	152	156					2	2	3	14	10	12	11		1	1	56	3	168	504	13	78		65		195
	147	151					2	3	7	5	12	9	8	3			49	2	98	196	19	66		47		94
	142	146					7	12	10	17	17	8	15	5	2		93	1	93	93	55	108		53		53
G	137	141		1		3	4	14	20	24	21	11	9	2		1	110	0								
	132	136		2		4	9	18	18	17	8	8	4	3	1	1	93	1	93	93	109	59	50			50
	127	131	1	1	1	7	7	11	15	16	18	9	5	2			93	2	186	372	104	59	45			90
	122	126		1	4	2	12	17	16	14	4	5			1		76	3	228	684	120	19	101			303
	117	121		2	2	10	9	6	6	7	2	2	2				48	4	192	768	107	12	95			380
	112	116			2	7	3	3	3				2				20	5	100	500	56	6	50			250
H	107	111		3	1	5	2	1	1								13	6	78	468	52		52			312
K	102	106			1			2	3	1	1						8	7	56	392	12	1	11			77
L	f		2	10	11	38	57	93	106	126	109	87	75	23	9	4										
M	x		7	6	5	4	3	2	1	0	1	2	3	4	5	6										
N	fx		14	60	55	152	171	186	106		109	174	225	92	45	24										
P	fx²		98	360	275	608	513	372	106		109	348	675	368	225	144										

$\Sigma x_{pos} = 669$	$\Sigma y_{pos} = 880$	$\Sigma xy_{pos} = 2783$
$\Sigma x_{neg} = 744$	$\Sigma y_{neg} = 933$	$\Sigma xy_{neg} = 0$
$\Sigma x = -75$	$\Sigma y = -53$	$\Sigma xy = +2783$
$\Sigma x^2 = 4201$	$\Sigma y^2 = 7377$	$N = 750$

$d_x = \frac{\Sigma x}{N} = \frac{-75}{750}$

$d_x = -.10 \quad d_x^2 = .01$

$d_y = \frac{\Sigma y}{N} = \frac{-53}{750}$

$d_y = -.071 \quad d_y^2 = .005$

$\sigma_x^2 = \frac{\Sigma x^2}{N} - d_x^2$

$= \frac{4201}{750} - .01$

$= 5.60 - .01 = 5.59$

$\sigma_x = 2.36$

$\sigma_y^2 = \frac{\Sigma y^2}{N} - d_y^2$

$= \frac{7377}{750} - .005$

$= 9.835 - .005 = 9.83$

$\sigma_y = 3.14$

$$r = \frac{\Sigma xy/N - d_x \cdot d_y}{\sigma_x \; \sigma_y}$$

$$= \frac{2783/750 - (-.10 \times -.07)}{2.36 \times 3.14}$$

$$= \frac{3.71 - .007}{2.36 \times 3.14}$$

$$\boxed{r = +.50}$$

Comments:

X-variable Height in inches
Y-variable Weight in pounds
Subjects: Students at Ohio State University
Computer: Th Date: January 17, 1922.

CORRELATION DATA SHEET
L.L. Thurstone
Division of Psychology
Carnegie Institute of Technology
Pittsburgh, Pa.
Form 1373

Editorials

THE PSYCHOLOGICAL DETERMINIST; OR DEMOCRACY AND THE I. Q.

At the recent convention of school superintendents at Chicago Dr. Bagley hurled another of those effective oratorical bombshells for which he is so justly famous. This one, like most of those which have preceded it in recent years, was an attack against certain educational trends which Dr. Bagley regards as inimical to the fundamental purpose of education as an instrument of democracy. This time, however, instead of repeating his customary frontal attack upon vocational education, Dr. Bagley executes a spectacular flank movement on intelligence testing, a development which he looks upon as an abetting ally of specialized training and as a challenge to those who believe in providing a liberal education for all the children of all the people.

The writer did not have the good fortune to hear Dr. Bagley's able address. He has only read it.[1] From an examination of the cold print he has not been able to see the justness of Dr. Bagley's position, though he would not regard it as strange if many who heard the address were convinced by its eloquence and clearly evident sincerity.

The keynote of Dr. Bagley's argument is that the psychologists who are promoting the intelligence testing movement are fatalists; that they exaggerate the factor of native endowment; that they are more interested in labeling stupidity than in curing it; that their doctrines lead inevitably to the neglect of average and inferior children and to an undemocratic emphasis on training the gifted; that their results are an excuse for limiting the education of the masses to narrow vocationalized courses; that the entire movement, founded upon little but psychological fallacies is subversive of American ideals of democracy.

---

[1] Bagley, William C., "Educational determinism; or democracy and the I. Q." *School and Society*, 15: 373, April 8, 1922.

The thing that seems to be especially repugnant to Dr. Bagley is the idea of predicting a child's development or drawing any conclusions regarding his native limitations from intelligence test data. He apparently prefers to base education upon a sublime faith that anything is possible for any child and that processes alone count. He seems to derive satisfaction from the fact that child nature contains so many unknown quantities. He would prefer to believe in "miracles" rather than search for the order and sequence which would make prediction possible. By implication he seems to deny any possibility of a science of education, for of course the essence of science is that it enables us to predict. If he is correct, then the educator is forever doomed to grope in darkness; there can be no light without the possibility of forecasting with some degree of assurance the limits of effect producible upon a given child by given educational processes.

Does Dr. Bagley really believe that native differences in intellectual endowment are mythical and that the product of the school can and should be made uniform? He is careful not to say that he believes this, but except in a single paragraph the implications of his address suggest that if he accepts inborn individual differences at all he regards them as of little if any significance. He feels that to acknowledge such differences would necessitate the abandonment of the democratic ideal, as he conceives this ideal. Therefore they are non-existent, or practically so. Thus we have a new and interesting application of Christian Science Psychology.

In the exceptional paragraph referred to, Dr. Bagley essays one timid glance at reality, but refuses to acknowledge the consequences of what he has seen. "Grant differences in native mentality? Certainly, such differences are inevitable." But almost immediately his vision is blurred by the moist tears of sentiment and the eloquent address closes with a rhapsodic peroration on the miracles that skillful teachers work with morons and on the ultimate illumination of the world by gleams of light struck from dull minds.

In another section of his address Dr. Bagley, referring to the supposed limits which native capacity places upon development, makes a great deal of the distinction between two kinds of mental growth,—the vertical and the horizontal. Vertical growth is improvement in the ability to perform more and more difficult types of intellectual operations. Horizontal growth is increase of one's repertoire of performances on a given intellectual level. Dr. Bagley admits that vertical growth has its (predetermined?) limits, but thinks he has won his argument by showing that the possibilities of horizontal growth are limitless. In making this admission Dr. Bagley has given away his whole case. Presumably he is not aware of the fact that this very distinction was drawn by Binet fifteen years ago and that it has been especially insisted upon by all psychologists in the field of individual differences. In reality it is one of the main arguments in favor of a differentiated curriculum. The differentiated curriculum allows all the vertical growth of which a particular child is capable and follows this with specialized training along horizontal lines. But hardly has the author admitted the limitations of vertical growth when, as if realizing that he had admitted too much, he retracts and argues that the common man in his daily business constantly grows in power to deal with abstract problems. "I assert most emphatically that he does."

Surely it is not necessary to remind the reader that assertions, however emphatic, do not constitute scientific evidence. The army mental tests give a good deal of opportunity for a subject to show how much vertical growth he has attained in ability to deal with the abstract, but careful analysis of army test results yielded no evidence whatever that any such vertical growth occurs between the ages of 21 and 31. If Dr. Bagley thinks that this negative result is due to the use of a faulty instrument it is to be hoped that he will devise another capable of yielding more valid data. But from his apparent faith in pronouncements based upon inner vision one wonders whether Dr.

Bagley would give much weight to any kind of objective evidence the experimentalist might be able to bring forward.

According to Dr. Bagley, the burden of proof is entirely upon the "determinist." Until he has demonstrated beyond the possibility of doubt that heredity sets its limits, it is our duty to educate all children as if such limits did not exist; until he has demonstrated that the low I. Q. can in no case be brought to average by the arts of pedagogy, it is our duty to act as if all of them could be. In thus resting his case upon the logical impossibility of proving a "universal negative," Dr. Bagley clears the road for the sentimentalist and removes every obstacle from the path of faith. But we can hardly admit the justness of this position. No one has ever proved the impossibility of life returning forty-eight hours after death, yet we do not hesitate to bury our dead within that time. No one ever has proved or ever will prove that the spirit of a departed soul may not return to earth, yet not many of us feel compelled for that reason to believe in ghosts. With respect to any issue we should be guided by the preponderance of evidence. From Galton on down to Thorndike and Davenport the scientist has produced a considerable amount of evidence in support of the hypothesis that one's mental traits, especially his intellectual abilities, are pretty largely determined by native endowment. The hypothesis may or may not be modified when all the facts are in, but by the canons of scientific method it is entitled to stand until concrete evidence has been adduced to show that it is false. Certainly it is entitled to the right of way as a guide to action until something more substantial than sentiment can be brought to bear against it.

The truth is, the spectacular fight which Dr. Bagley wages is largely against a straw man. He reads a meaning into the words of the "determinist" which was never intended and accepts as just the interpretations placed upon their findings by the sensational paragrapher. The psychologist does not propose, as Dr. Bagley unjustly asserts, "to exclude everyone except the high I. Q.'s" from the privileges of secondary education, though he pleads guilty of demanding new types of secondary education which would be better suited to inferior intellects than is the typical college preparatory curriculum. The present writer thinks he is fairly well-acquainted with the educational views of his co-workers in the mental-test field, and he has yet to hear from any of them the suggestion that the years of schooling should be reduced for dull children. On the contrary, their constant argument has been that the present shameful record of early eliminations is largely due to an ostrich-like unwillingness to face the facts of individual differences; that a differentiated course of study based upon a frank recognition of these differences would keep the low I. Q.'s in school and insure that their time there would not be wasted. Nor does the psychologist advocate the kind of vocational education for low I. Q.'s that would make them mere adjuncts to machines in the industrial world. He believes, on the contrary, that it is the prescription of a uniform curriculum for all which is most likely to foredoom the dullard to the role of an industrial automaton.

The writer has been considerably puzzled in trying to interpret Dr. Bagley's flings at "social leaders" and the "aristocracy of brains." "It would be easy to show that an intellectual aristocracy is just as reprehensible as an aristocracy based upon family or wealth . . . ." Social leaders, he contends, do not need to be trained; all that is necessary is to train the common man so that he will know when to tell his leaders "where to get off." Special educational opportunities for gifted children are contrary to the spirit of democracy. Presumably, also, it must be undemocratic to provide expensive colleges and universities which only the superior youth may enter. It would be most undemocratic to remind a class of college graduates that they are marked by ability and training for leadership and that they should be sensitive to the obligations which this fact imposes. It is sufficient to teach the common man to tell his would-be leaders "where to get off!"

One can only surmise what Dr. Bagley thinks his opponents mean in urging more attention to the training of social leaders. By social "leadership" he seems to understand some kind of political or industrial "drivership." He forgets that social leaders include also the Luthers, the Newtons, the Shakespeares, the Darwins, the Pasteurs, the Blackstones, the Edisons, the President Eliots, the Curies, and the Herbert Hoovers. Studies of the biographies of eminent men and women so often reveal the role played by happy accident in bringing the necessary opportunity for genius to "out," that one can be reasonably certain our ancient policy of "laissez faire" has lost us many a potential leader in the realm of science, literature, politics, and morals. It is entirely improbable that the name of Newton would have been known today but for the fact that he had a discerning uncle who recognized the boy's intellectual superiority and had him sent to a school for gifted youth (Cambridge). Thus by sheer accident the greatest intellect of modern science found its opportunity.

Far from believing (as Dr. Bagley charges) that the function of the school is to certify brain quality rather than to teach, and far from encouraging teachers to neglect the dull as unworthy of their efforts, the psychologist of individual differences believes that the one purpose of intelligence tests in the schools is to aid us in making the most of every child, the dull as well as the bright. Anyone is at liberty to disagree with him as to the fitness of the method for this purpose, but the purpose itself ought not to be misconstrued.

If anything has been demonstrated by psychology in the last quarter century it is the existence of native differences in mentality. If the psychologist can escape the charge of being an enemy of democracy only by denying his facts, then he must admit that he is an enemy of democracy. The writer, for one, would insist upon a less naive definition of the term democracy. Certainly any definition to be acceptable will have to square with the demonstrable facts of biological and psychological science.

LEWIS M. TERMAN

Leland Stanford University

RETRENCHMENT

Wherever one goes and whatever meetings one attends, one finds that the burden of the anxiety of superintendents and other administrative officers centers around the fact of reduced revenue. In many localities valuations have shrunk, and everywhere the general business depression has tended to reduce rates. Thus from coast to coast those in charge of schools are desperately figuring how they can make their expenditures come within a diminished budget.

It would be remarkably instructive if we could collect the ways in which retrenchment is being effected in various city systems. Action in this respect, as in so many others in life, depends upon our ideas of the value of rival procedures, institutions, or acts. When you tell a friend you haven't had time to write to him, you do him deep dishonor because your days have been full of a number of things, including some of little consequence; yet you have considered them all more valuable than the writing of a letter to him.

Not widely different is it in operating a school system. The things we insist upon maintaining are the things we think most valuable; and the things we first dispense with are those we consider least essential. No matter what fine sentiments we may have expressed in other days, when we are confronted by the cold, hard alternative of making a choice, we come surprisingly close to an expression of our real sentiments.

A superintendent the other day was rather complacently announcing the way in which he had met the problem in his city. In order to maintain his salary schedule (which was one of the things he thought of high value), he had decided to decrease the number of positions by increasing the number of pupils per class. He found comfort in contemplating the investigations of such men as Cornman, Boyer, Harlan, and Elliott, who had shown that almost, if not quite, as effective teaching could be done in large classes as in small—provided the classes weren't *too* large. What seemed to annoy him most was that he had for years been advocating small classes and had succeeded in reducing their size because of his persistent advocacy. Now, however, in deciding upon this particular type of retrenchment he was placed in the uncomfortable position, as he said, of having to "eat crow." Indeed, he seemed somewhat dazed to find himself maneuvered by the logic of events into his present position. Evidently, it was a surprise to him to find that his advocacy of the small class was really a matter of theory, and that he was willing to abandon his notions when confronted by a condition rather than a theory. The acid test of a money shortage had brought to his attention the fact that he really did not set a high valuation upon the small class.

Similarly, other superintendents are finding that some of the things they have advocated have less real value than they have hitherto been trying to induce others as well as themselves to believe. For example, one superintendent says in substance: "What is industrial education in the senior high school for? I have an expensive equipment and personnel in my senior high school for the purpose of teaching printing, sheet metal work, pattern making, etc. Now I know I am not going to train these children to be printers, or sheet metal workers, or pattern makers. Why am I offering these courses?" This is a question which might well have been asked a few years ago when the money was first appropriated for the installation of these courses. They are probably just as valuable now as they ever were.

Other superintendents are reducing auxiliary service—such as evening schools, play grounds, community centers, and lectures. Americanization, about which we have heard so much, is being curtailed in some quarters. Other superintendents are questioning the value of the kindergarten. One says: "The kindergarten people have never proved their case. They have never even tried to. With instruments of measurement lying all about them for the past few years, they have continued to sustain their cause by a sort of mid-Victorian sentimentality. It looks to me as if the kindergarten were a luxury."

So one after another of those who control the destinies of the public schools decides each in his own way what he thinks is least valuable among the educational activities of his community. There is very little that is scientific about this. The bases on which the curriculum is simplified, special classes for gifted children are discontinued, and summer classes are abolished, are no more respectable than were the bases on which these activities were established. The practical school man is curiously situated. He must make decisions without having much evidence to go by. Educational research lags so far behind educational practice that it lights the way long after many a traveler has passed by and either lost his way or reached the wrong destination. There is, as we have before said, a need of quantity production in research. By the way, we have not heard it said—although it is doubtless true in some quarters—that retrenchment has led anyone to abolish a department of research where such a department has been established.

On the whole we believe that the necessity for retrenchment is producing more high-class thinking among school men than they have been accustomed to for a long time. Despite the fact that decisions must often be made hastily, we suggest that the superintendent at least owes it to his community and himself to familiarize himself with the evidence concerning matters which he contemplates dispenses with or curtailing. Any good bureau of reference and research can give him some guidance in this particular. But in the name of a long-suffering school public, don't let him broadcast a questionnaire!

With all this effort at retrenchment, however, the better superintendents are not advocating the reduction of teachers' salaries. There is a disposition to maintain the present schedule at all events and at all hazards. We believe this policy is wise; and we advise teachers who may find themselves called upon to teach more pupils or otherwise to increase their teaching load to think twice before criticizing the administration. What the administration needs and ought to have in its efforts to protect salaries is the enthusiastic support of the teaching body. There is a real, and in many cases a desperate, shortage of funds. Information on this point should be honestly and completely presented to the teachers so that they will understand this to be something more than the annual palaver about high taxes. If this is done, teachers should be expected—indeed they themselves should expect—to assume whatever additional burden may be necessary to save the day for teachers' salaries. In truth nothing but a short-sighted policy would induce them to do anything else. Even intelligent self-interest should lead them to do more work in the next year or two in order that the pay and the respect of teachers may suffer no impairment.

B. R. B.

Reviews and Abstracts

JORDAN, RIVERDA HARDING. *Nationality and school progress; A Study in Americanization.* Bloomington: Public School Publishing Company, 1921. 105 pp.

Dr. Jordan chooses Minneapolis and St. Paul as his field for study. His investigation is a very thorough one and certainly no time nor effort has been spared in trying to determine whether the factor of nationality of school children enters into their progress in school, in so far as this can be determined by a comparison of children of foreign born ancestry and those of distinctly American parentage. The study, however, is too scientific to make it of real value to either the teacher of foreigners or to superintendents. Dr. Jordan attempts to refute the popular phrases of the "bright Jewish children," "stolid Scandinavians," "mercurial Irish," etc.

Studies were made investigating the birthplace of parents, grandparents, and children, also as to whether English was spoken in the home; factors of retardation, acceleration, and school marks were studied. Objective mental tests were given; and other studies were made effecting school progress along such lines as mobility of students, occupation and economic status of parents, home conditions, church attendance, etc. From all these studies involving 39 printed tables, Dr. Jordan comes to the following conclusions:

1. There is no justification either to despise the attainments of the foreign born and his offspring or to sublimate his achievements unduly at the expense of the native American.
2. There are no such marked differences as to justify the popular phrases previously cited.
3. The handicap of the language difficulty does exist and it should be the concern of every American to see that all possible steps are taken to remove this difficulty.

4. There must be more flexibility of the school curriculum in order to reach special foreign groups.

Dr. Jordan's feeling relative to the imperative need of a change in the present Naturalization law will hardly meet with the approval of those who know and work with these foreigners. His judgment is that an applicant is not fit for Citizenship unless his family, as well as himself, speak English, unless he reads only an English newspaper, and unless he attends a church where only English is spoken. It certainly is not true that it is impossible for a man to be a good American citizen until he withdraws himself from any contact with his former homeland. It is quite possible that through these very same contacts he may be of greater influence to his friends of foreign birth. The women in the home are especially handicapped because of the lack of time and opportunity to acquire the English language and it would seem hardly fair to keep men from citizenship because of this almost unsurmountable difficulty.

Springfield, Massachusetts JOSEPHINE D. MASON

ANDREWS, FANNIE FERN (Secretary). *An American citizenship course in United States history.* New York: Charles Scribner's Sons, 1921. 5 volumes.

In 1913 a committee of The American School Citizenship League began working on the report which has recently appeared under the above title. This committee was composed of Messrs. W. F. Gordy, P. P. Claxton, C. E. Chadsey, J. H. Van Sickle, John H. Hall, Mrs. J. W. Hall and the Secretary, Mrs. Andrews. After sitting in "earnest session as many as twenty-six days" the committee was able to lay the foundations for the rather voluminous report which is now before the teaching public for acceptance or rejection.

The course proposed is much like the one suggested by *The Committee of Eight* in 1909. Inasmuch as this report is so well known it seems a waste of valuable space to describe the one under review in any detail. It should be said, however, that instead of stopping after making a few general suggestions as *The Committee of Eight* did, this committee compiled a considerable body of type-study material. This material appears in four books, one for Grades IV and V and one each for Grades VI, VII and VIII. Since the history recommended by the committee for Grade VI, VII and VIII is so universally taught at the present time in these grades this type-study material ought to be of considerable assistance to upper-elementary grade history teachers.

Generally speaking the report will disappoint the reform element in our midst. Many will designate it backward rather than forward looking. It may be we have made no progress since 1909 in the matter of elementary-grade history. Evidently the committee feels that we have made but little, if any. One must not be too critical, however, of the results of the Committee's work, for two of its members, Messrs. Gordy and Van Sickle, were members of *The Committee of Eight* and one would not expect them to admit that there were any serious defects in that committee's report. R. M. TRYON

University of Chicago

HOSIAC, JAMES FLEMING. *Empirical studies in school reading.* (Teachers College, Columbia University, Contributions to Education No. 114) New York: Teachers College, Columbia University, 1921. 174 pp.

The avowed purpose in making this study is "to establish a method of evaluating (literary) reading-books intended for use in Grades IV to VIII of the elementary schools in terms of the questions and other helps to study which they contain." The investigation is carried out along four main lines; (1) The opinions of authorities as to the aims and methods of teaching literature in school; (2) the analysis and classification of the *study helps* found in four sets of popular literary readers; (3) evidence as to the common practice of teachers in teaching literature; (4) experimental teaching to determine the relative value of methods and devices.

The principal conclusions reached in each line of investigation follow: (1) "The weight of authority concerning the teaching of literature to children appears to be in favor of informal methods of treatment—looking to appreciation or enjoyment," and that the recitation "should be devoted chiefly to hearing the selection well read and seeking to enter into the experience which it provides by the exercise of imaginative sympathy." (2) In the four sets of readers "the proportion of space devoted to the notes and helps varies from 15 to 38 percent." The total number of questions was 18,873. The average number per volume was 943. "The overwhelming majority of these questions were found to relate to details, particularly linguistic and informational details." (3) From a study of stenographic reports of eighteen lessons in literature by eighteen different teachers it was concluded; (a) That the teachers emphasized the same interests that were emphasized in the books. (b) That they make little use of questions and study helps, preferring "to conduct their reading lessons in their own way without reference to studies found in textbooks." (c) That there was "a tendency in practice to deal with details at the expense of the whole." (4) "Evidence was obtained sufficient to indicate a strong probability that a piecemeal attack on poetry with formal and ineffective reading aloud and much insistence upon the discussion of details of language, without any attempt to bring the pupils' own experience to bear on the poem as a whole and without any final synthesis, is relatively inefficient as a method of teaching." Hence, "the emphasis was wrongly placed by the editors of the textbooks examined and by the majority of teachers whose lessons were reported."

The important conclusion seems to be that literary reading books are to be valued according to the amount of good literature they contain in proportion to notes, questions, and other study helps. And since such study helps are not in accord with the accepted aims of teaching literature, and since teachers make little use of them anyway, the "space devoted to such studies" is mostly wasted and "might well be given up to good poetry and good prose."

The author's conclusions, so far as they pertain to literary reading-books and to the teaching of literature, are probably valid. But when he confuses the teaching of literature with the teaching of reading, there is room for difference of opinion. His idea seems to be summed up in a quotation from Kendall and Mirick, "In the reading lesson, let the aim be to awaken in them (the pupils) a genuine love of reading and to lead them into the broad fields of good literature."

Some of us believe that the real aim in the intermediate grades should be to teach children how to read. Only so can they read comprehendingly enough to enjoy good literature. Evidence as to method and content as brought forth by the author indicate the wisdom of divorcing the teaching of reading from the teaching of literature and suggest that purely literary material is not the best material to use in teaching reading.

Indeed it is suggested that the discrepancy between theory and practice which the author finds may be due to the fact that the selections were beyond the comprehension of the pupils, thus necessitating much analysis and questioning upon details.

As a treatise on the materials and methods of teaching literature, the book deserves careful reading. It is well planned and well written, and it contains many valuable suggestions.

S. S. Brooks

Winchester, New Hampshire

Dunn, Fannie Wyche. *Interest factors in primary reading material.* (Teachers College Contributions to Education No. 113) New York: Teachers College, Columbia University, 1921. pp. 70.

The complex field of reading needs scientific evaluation of subject matter and method of procedure. Miss Dunn's contribution is notable in that she attacks the key problem involved—children's interests in reading—and in that she recognizes the importance of these interests in the primary grades, beginning with the first.

After summarizing the interest appeals in reading material of the last one hundred years, two direct challenges are made: that "what is of interest to children of this age has never been demonstrated" and that "no justifiable criterion of literary merit has been set up in making these selections." Yet for a century we have talked of interest as a factor and for half a century have stressed literary values. Fortunately Miss Dunn prophecies that exact scientific information will come in this field as it has in others, then makes her own contribution.

Thirty-one selections comprising fiction, factual material, and poems were selected by competent judges from 243 samples of primary reading material. These selections were read in pairs to 195 different classes in grades one to three, distributed over five states and including a wide range of child types. The children voted on their choice of each pair of selections. Their catholicity of taste is in sharp contrast to current opinions as to their favoring fancy. Poetry receives a very decided blow, indicating that far too much of it is included in our texts.

Twenty qualities present in high degree in the samples were selected by judges, and correlations worked out to show their influence. Some unexpected results appear here. Surprise ranks highest with "animalness" about equal to it for boys, and childness and familiar experience for girls. Plot, narrativeness, liveliness, conversation, and moralness rank high, while humor and familiar experience are much less important than one would expect. Slight sex differences appear.

The rating of selections according to adult judgments proved little more valuable than a random selection would be. A few individuals succeeded fairly well in making such a rating.

The study closes with a plea for further studies in this field, for an extension of the types of material in our texts to include all dawning interests and to awaken those not aroused, and for the selection of a core of minimum essentials to be given to every child. Especial attention is called to the need for factual material of better quality.

Alert primary teachers will find suggestions here for original work in preparing and testing out new types of material for their grades. Teachers of method courses will appreciate the text for class use. Students of education may gain from it courage to attack vital problems rather than the formal habit-forming procedures which allure because they are so easily reduced to statistical form. Child study specialists will rejoice that a way has been discovered in this important subject for children really to speak for themselves.

FRANCES JENKINS

University of Cincinnati

MYERSON, ABRAHAM. *The foundations of personality.* Boston: Little, Brown & Company, 1921. 406 pp.

Dr. Myerson's book is another venture in the field which John Stuart Mill denominated ethology, and regarded as one of the most important of the sciences. Ethology has in recent years received many interesting contributions, each of which has its distinctive value. Shand's "The Foundations of Character" describes the history of character more from the point of view of the traditional introspective psychology. Jastrow's "Character and Temperament" is popular in presentation, and adds the angle of vision of the experimentalist and the student of unusual types. Wells's "Mental Adjustment" contributes the insight of the psychologist having an extended experience with insanity and various other forms of psychic abnormality. Hocking in his "Human Nature and Its Remaking" gives us the attitude of the philosopher and especially strives to show the significance of religion in the remaking of the natural man.

Dr. Myerson is a physician who specializes in psychiatry and whose experience brings him in contact with many cases varying not widely from normality. He abhors metaphysics and writes in a very interesting style, giving us a popular account of the growth of personal character, its types and the influences that play upon it leading it to attain various kinds of efficiency or to become abnormal or perverted. His account begins at the bottom with the organic and environmental bases of character. It sketches the elementary psychology of memory, habit, emotion, instinct, intelligence, purpose, and will. The influence of various emotions and emotional attitudes in releasing energies and in providing compensations is extensively treated. The theories of Freud are recognized, but the author feels that they afford a very one-sided interpretation of the conflicts in character. The point of view is eminently sane and well-balanced. While a large part of what he says is just common sense in the interpretation of human nature, he supports his views by a wide range of facts drawn from his medical training and experience, and from an extensive acquaintance with psychology, philosophy, and literature. Not a little of the interest that he evokes is due to his ability to call to his help Dickens, Moliere, or Goethe; Lecky, James, or Bergson, as well as the authorities of physiology or psychiatry and the like. His book is one that may be read with profit and interest by teachers and general readers, and the professional psychologist will not only approve of its sanity, but will find in it many valuable suggestions.

ERNEST N. HENDERSON

Adelphi College

BODE, BOYD H. *Fundamentals of education.* New York: Macmillan Company, 1921. pp. xi + 245.

We have been so busy in these recent years investigating and checking up results in the field of education that we have not had the time or the disposition to take our bearings and find out whither we are going and why. With the exception of the notable writings of John Dewey the speculative aspects of education have been conspicuous by their absence. So we can heartily welcome this new book by Professor Bode in a field new to him. The author has already obtained an enviable reputation in pure (?) philosophy. As a logician and epistemologist of the pragmatic school his writings are well and favorably known. This excursion into the philosophy of education, which by the way promises to be more of a permanent adventure than a brief foray, is full of promise. This first book in educational philosophy is not the last that Dr. Bode will write. I predict that it will be followed by others of perhaps a more fundamental and penetrating nature, but by none more entertaining and vital.

In this essay the writer runs true to form as a pragmatist. He tells us that education must have a definite sense of direction. "Unless we know where we are going there is not much comfort in being assured that we are on the way and traveling fast." But just where are we going; what should be our destination? We look for an ultimate educational aim, but Professor Bode adroitly avoids giving us this information. All aims are good as long as they are relative. There is no *summum bonum.* The true aim of education is to foster intellectual and spiritual growth. "We cannot solve the problems of life in advance. Aims must be worked out as we go on." This means that the most important factor in education is the teacher who is to create and foster the spirit of open-minded inquiry. Thus Bode,—but does he then mean that the end of education is the search for truth? Yes, and no; for while he implies this he asserts more. It is not any kind of growth, not any sort of intellectual activity that will serve the purpose. The growth must be in the direction of social understanding and sympathy. It must result in an evaluation of human life, and teaching must be in terms of a social context. We must arrive at a greater cooperation between members of a social group and also between different social groups. We must realize democracy more completely,—not merely political but social as well. Thus the aim of education is growth in the direction of greater human understanding and cooperation.

But is not this setting up an ultimate end? It may seem so but Professor Bode is not to fall into such a trap. He tells us that *social* ideals must be developed, but that these ideals are not to be final; they are merely stepping stones. To know the final aim of education would be to know human life completely as it was, as it is, as it shall be. Since this is impossible, for life is as it is lived, and the end is never absolutely achieved until life is at an end, we can have aims only that are relative and tentative. Although Bode does not say this, it is to be inferred that any scheme of social order that we accept to-day may be discarded later. It is also possible that we may be going but not in the right direction. The best that educational philosophy of this type can offer us is to indicate the progress that we are making along a road that leads somewhere. But is this the road we should take if we wish finally to arrive? Probably not, because we shall never *finally* arrive. As long as the road is pleasant and our companions agreeable, and we have plenty to do to keep us alert and active, what more can we expect? So it seems to me we must confess that we are on our way and traveling fast in a certain direction that surely is not a final stopping place. If we keep on we will arrive, put up over night, and in the morning start our car, step on the gas and speed along following the next guide post and reaching the next town, and so on just forever and ever.

Pragmatic philosophy can say nothing more. However, to my way of thinking this relativity is better than the absolutism of the past. Education has suffered far too much because standards have been set up that are unvarying and eternal throughout the ages. In Dr. Bode's book there is life, movement, adventure, the breath of the open country, and the lure of the distant horizon. The discourse is pleasant, well-written, entertaining, replete with common sense, and above all it stands the final test of the pragmatist,—*it works.*

STEPHEN S. COLVIN

Brown University

BROOKS, FOWLER DELL. *Changes in mental traits with age determined by annual re-tests.* (Columbia Contributions to Education, No. 116.) New York: Teachers College, Columbia University, 1921. 86 pp.

The great desirability of repeated measurements on the same individuals over a term of years has been frequently pointed out, not only for the light it throws on changes in mental traits with age, but also on many other related problems. This study, based on re-tests of 171 pupils from the fourth to the ninth grade, ranging in age from nine to fifteen years, is a contribution to this problem. Twenty-five mental and educational tests were given three times, giving a record of change over a two-year period, though only sixty-seven of the subjects took all the tests three times.

An historical survey of the more significant of the many studies on changes in mental traits with age shows few investigations in which the same individuals have been re-tested, and these are chiefly Binet re-tests to determine the constancy of intelligence quotients.

After reporting in detail gross scores and the median yearly improvement for each age, test, and sex, the writer sets out making the different gains and rates of gain comparable by dividing the different gains by a constant—the average of the eleven-, twelve-, and thirteen-year standard deviations of each test for each sex. The resulting table shows a bewildering variation in the amount of gain in the different tests at different ages. They resisted interpretation without classification. The data were classified on a basis of similarity of functions (simpler, memory, higher or more complex, and informational), presence or absence of high scores, influence of school instruction, and ability to make initial scores. When allowance is made for selection which raises the scores at the earlier ages and lowers them at the later ages, and the curves are smoothed by increasing the number of cases and the number of tests, the conclusion is reached that not only are the gains positive at all ages between nine and fifteen for both sexes, but also the rate of improvement is probably best represented by a straight line whose slope is probably slightly negative. A comparison of rates of gain in other investigations, in which case the changes were determined by taking the differences between the central tendencies of groups of different ages, shows substantially the same tendencies except that the gains are much less.

There is no evidence in these data of periodicity or rhythm in mental growth, no evidence of pre-adolescent or adolescent spurts; no evidence of a slowing up in rate after eleven or twelve; and no evidence of a cessation of mental growth before fifteen, the highest age tested.

The correlations between mental traits measured at a two-year interval for the four types of functions are high—the corrected coefficients range from .64 to .99—indicating that the abilities are a relatively permanent endowment. The correlations between gains in different groups of functions are positive though low, and indicate no evidence for the doctrine of compensation, that the gain or improvement in one group of functions stands in inverse relationship with gain in other functions.

This investigation represents the sort of study we greatly need, though the difficulties of giving tests at more frequent intervals than a year without having the results as measures of mental growth vitiated by practice, and over a longer period than two years without being vitiated by selection or reduction to too small a number of cases to be significant, are very great. Accurate individual growth curves, however, we must have, and the sooner we set about getting them the better off we shall be.

V. A. C. Henmon

University of Wisconsin

Ensign, Forest Chester. *Compulsory school attendance and child labor.* Iowa City: The Athens Press, 1921. 263 pp.

One scarcely expects to find more than a recital of facts and figures of a somewhat prosaic nature in a treatise on the above topic. Consequently, the fact that the author has here presented a very thorough study of compulsory school-attendance and child-labor in a manner so vitally interesting, that the reader's attention is held at every point, is no mean compliment to him. Were there more presentations of this subject written in such a simple yet convincing manner, we might hope to gain more careful consideration of this important question.

Without undue and tiresome detail the author outlines the status of education in relation to employment in England. From this he passes to the development of compulsory education and compulsory employment in America during the colonial and the early national periods. Having laid this broad foundation for the reader, he undertakes to show the gradual development of the movement to protect the child's right to gain an education and in his right to live a normal child-life instead of one of drudgery and practical slavery under unscrupulous masters in industry. Five states, Massachusetts, Connecticut, New York, Pennsylvania, and Wisconsin are chosen to demonstrate the principles which have been operative and those which should be encouraged or avoided, depending upon the desirability of education or the undesirability of child employment. Other states notable in industry and in education were necessarily eliminated because of the incomplete and hazy records of attendance,—a serious indictment against the educational officers of those states.

The purpose of the book stated in the introduction is "to trace in some detail the development of legislation for the control and compulsory education of children from its inception in the English statutes for the restraint and industrial training of the children of the poor to its expression in the elaborate systems of universal, obligatory education with the accompanying elimination of child labor in the most progressive American states." This line of thought is carried throughout the study, giving it a unity and conciseness seldom found. Furthermore, each chapter is carefully organized as a distinct unit in itself with definite purposes and conclusions. An extensive bibliography is appended which is of value not only in giving the authorities consulted, but in furnishing an excellent guide to the student in this field.

C. C. McCracken

Ohio State University

Toops, Herbert Anderson. *Trade tests in education.* (Teachers College Contributions to Education No. 115.) New York: Teachers College, Columbia University, 1921. 118 pp.

Mr. Toops enjoyed an unusually varied experience in the making of trade tests for army use. The book under review is an attempt to summarize the best of his army experience for the use of educators in both academic and vocational schools. The book is full of suggestions for any one who is interested in improving methods of evaluating the results of teaching.

The terms "trade tests" and "vocational tests" are used synonomously by the author who points out that these differ from other tests primarily in their content. He gives the following classification:

I. Tests of proficiency or success in a trade.
    1. Verbal.
        a. Oral administration.
            (1) Army oral.
                (a) Oral answers.
                (b) Written answers.

(2) Pictures.
b. Written administration.
(1) One-word answer.
(2) Multiple choice.
2. Performance tests.

II. Tests of trade capacity, potentiality, or promise. (This group includes practically every known variety of test, including the above verbal forms.)

The author illustrates the several types of trade tests; points out their usefulness as a basis for improvement of instruction, and as a means of predicting success in employment; gives a good analysis of how to make a trade test including the assemblying, revising of material, and the standardizing in terms of groups of known ability; discusses methods of administering tests and of safeguarding teaching and coaching; considers the adaptation of trade-test technique to school examinations; and closes with a chapter of "general considerations in testing and training for proficiency and promise in the trades." Several characteristic types of tests are described in detail.

Mr. Toops realizes the need for and the difficulties involved in making a thorough analysis of the requirements of the job before undertaking to make a trade test. He offers one concrete suggestion as to how to make such an analysis in the form of a "general job analysis questionnaire sheet."

His evaluation of the vocational guidance tests used in the Manhattan Trade School for Girls, New York City, indicates how such an experiment might be better organized, administered, and finally evaluated. He believes this experiment to be "among the foremost vocational guidance programs to be found today."

Students of statistical methods will find the volume worth their attention. Those interested in the philosophical background of trade education and trade testing will find that the author has positive ideas underlying his scientific attitude. The selected bibliography concerning trade tests is quite complete.

Trade Test in Education constitutes a first volume for everyone interested in trade testing. It will be particularly useful to those interested in making trade tests; and to teachers and administrators who are looking for methods of improving the measurement of teaching in their own schools—vocational and academic.

J. CAYCE MORRISON

The University of the State of New York

SMITH, EHRLICH E. *Teaching geography by problems.* New York: Doubleday, Page & Company, 1921. 306 pp.

There is an awakened popular and educational interest in the study of geography. The progress of the World War led the intelligent reader to study his atlas, and the maps in newspapers and magazines with unusual interest. An understanding of the consequences of the War required still keener appreciation of geographic relationships.

Educational journals have published articles concerning the necessity of making geography serve the present-day needs of our national life and its international relationships. Superintendents of schools have formulated courses of study in geography to make definite use of the awakened interest in the subject. New books on the teaching of geography have been published to aid teachers and pupils to use to advantage the present-day interest in geography.

Teaching Geography by Problems gives the teacher two valuable points of view concerning geography: (1) the importance of geography in the American educational systems of today; (2) the use of problems, or the problem-method, in teaching geography.

Teachers who have been requested to present their geography work by means of problems without knowing just what the request called for, will welcome this book for the clear, definite presentation of the meaning of problems, and the comprehensive development of selected problems. One chapter is given to a discussion of the following topics: how a problem is selected, organization of topics, quantitative and qualitative treatment, development of problem types, testing the problem, values attributed to the problem method, illustrations of problem types, minimum requirements of place geography.

In the development of these topics, a number of specific problems are stated, some of which are the following:

> Why have the Philippine Islands a promising future?
> Do you think the United States could afford to sell the Philippines?
> What has England done for India?
> Why are the Hawaiian Islands a valuable possession of the United States?
> Why has Great Britain the largest merchant marine and navy in the world? A suggestive topical outline for the development of this problem is given.
> Can Japan be to the East what the British Isles are to the West? This problem is so outlined as to use the knowledge gained in the study of the preceding problem.

Nearly one-half of the book is devoted to "Illustrative Problems" suitable for use in various grades.

A study of *Teaching Geography by Problems* will show the teacher that the "problem method" which has sometimes startled her is not something new and strange, but that much of her best teaching has always been done by the use of problems. The superintendent who wishes to develop the problem method in his schools in a natural and helpful way will find this book an aid to him and his teachers, not only in geography, but in other subjects as well.

DOUGLAS C. RIDGLEY

Clark University

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors. Address all correspondence concerning this department to Doctor E. J. Ashbaugh, Ohio State University, Columbus, Ohio.

"In the April number of the Journal (page 346) announcement was made of the Douglas Standard Test and the Gregory-Spencer Geography Test. The place of publication of these tests was erroneously given as the University of Washington. These tests are published by The Bureau of Educational Research, University of Oregon, Eugene, Oregon. We regret this error."

In an address telling of the uses of intelligence tests in Normal Schools, Miss Bessie L. Gambrill, affirms that intelligence tests will not give all the "facts that are required for prognosis of a student's probable success as a teacher. There are other qualities, such as the dynamic personality traits, the absence of which will cause failure in teaching, no matter how highly endowed intellectually an individual may be."

Mr. Edward D. Randolph, now assistant to Dr. Bagley at Teachers College, Columbia University, has been elected as professor of educational sociology at the University of Washington. He will enter upon his duties in the autumn quarter, 1922. Mr. Randolph was for several years in charge of the work in educational sociology at the State Teachers' College, Greeley, Colorado.

Superintendent Allen of Springfield, Illinois, gives research men a valuable suggestion when he proposes as an important practical school problem, the determination of the work of each grade, and in each subject; this work to be expressed in terms of achievement. Not only the teachers and school officials but interested parents would be benefited in their efforts to adjust the school to the child by knowing precisely what is expected in arithmetic in the third grade or in geography in the sixth grade, or in any subject and grade of the elementary school. Each school system should be prepared to furnish this information. Since there is no reason why the standards should differ, for the fundamental subjects at least, in the first six grades, we ought to have standards substantially uniform for all systems.

Superintendent Newlon of Denver reports that he has received 73 questionnaires since last December. Everybody, including Mr. Newlon, agrees that in respect to questionnaires "Something ought to be done." The questionnaire is a necessity. It is the abuse of it to which we are calling attention. If we had a coordinating agency which should advise makers of question blanks so that duplication might be avoided and frequent topics surpressed, an undoubted benefit would be conferred, not only upon those who receive such blanks but also upon educational research.

It is reported that Detroit has raised the number of high school pupils per teacher from 25 to 35. This has been done in order to reduce the number of positions and thus maintain individual salaries. What effect this will have upon the efficiency of instruction is not known. It is certain that many cities are confronted with reduced revenues in proportion to the work to be done, and Detroit is not alone in the adoption of this device.

We are informed that Teachers' College, Columbia University, requires graduate students and faculty members to secure approval of questionnaires. This statement of approval is put on the questionnaire where the recipient can see it. Either on the questionnaire or in the letter of transmittal, the request is made that questionnaires emanating from Teachers' College but not regularly approved be disregarded, and that if questionnaires without approval are received, Teachers' College be notified.

The Woman's College of Alabama has this year enlarged its Department of Education by the addition of a full course in practice teaching. This has been accomplished by cooperation with a local public school. The work was entered into with fine spirit on the part of the students and has proved most successful. A similar arrangement will be continued next year.

The Bureau of Education, Reference and Research of the University of Michigan is directing a cooperative testing program in that State, covering the fields of spelling, reading, and arithmetic in the elementary grades; and Holtz's Algebra, Brigg's English, and Henmon's Latin and French Tests in high school.

The spelling survey has one unique idea in that a uniform list selected from the Buckingham Extension of the Ayres Spelling Scale is to be used by all systems cooperating and an additional list made from the same sources by the superintendents involving words which have been taught sometime in the school life of the child will also be used. This will afford a check against the results obtained from the uniform list by using additional words which it is certain the children have been taught.

Some interesting facts regarding the health of gifted children has been revealed by a three-year nutritional experiment carried on by a New York City school.

"Gain in fitness was greatest on the part of unusually 'bright' children. A number of children having intelligence quotients of 125 or over were examined and a larger percent of underweight was found among them than in any of the groups studied, yet under improved conditions, they responded better than the others. 'Gifted' children probably need a more ample allowance of sleep, fresh air, and food than average children and are the very ones least likely to spend their time in out-of-door play or to go to bed early, with consequent lessened appetite and powers of assimilation."

As an evidence of the economy in administration which the administration of senior-high schools at Columbus, Ohio, has been putting into effect a showing is made of a reduction in the number of vacant rooms in the five high schools of the city together with a reduction in the number of teachers. The latter has been made in the face of an increase both in the number of classes and in the number of pupils. Rooms were vacant for 94 fewer periods per week this term than last, four fewer teachers were in service, 98 more classes were organized and the enrollment increased 451. The average number of classes per teacher last semester was 5.22, this semester 5.89. The average number of pupils per teacher was 24.3 last semester and 27.4 this semester. Anyone who knows anything about the administration of a large city system knows that these results were not obtained without close and careful planning.

Bulletin No. 1 of the Bureau of Business Research of the University of Illinois is entitled "Illinois Taxes in 1921." Certain pertinent facts, which are pointed out to the taxpayers of the state, are of particular interest to school people.

Nine-tenths of the money arising from the general property tax was spent for local purposes, and less than one-tenth was appropriated by the legislature. More than one-half of the money appropriated by the legislature was for highways and waterways; only one-sixth of the total amount appropriated by the general assembly was for education. The state has a large source of income other than property taxes, and this source is tapped to care for general expenses, while the school distributive fund is taken entirely from the general property tax. The average taxpayer on scanning his tax receipt gains the impression that increased appropriations for education are responsible for the increase in the state tax rate. This is decidedly incorrect and misleading.

Of all funds going to the local government in 1921, 37½ percent was expended for public schools.

Have you seen the new "School Accounting Report" developed and proposed by a joint committee representing the Bureau of Education, the Department of Superintendence of the N. E. A., The National Association of Public School Business Officials, and The Inter-City Conference? We have lately received a copy of Part II of this report. It has reference to financial statistics. Its classifications have probably been more carefully worked out than any classifications which have hitherto been proposed. The discussions and definitions which accompany the report add very greatly to its value, since they increase the probability that the reports will be made out in a substantially uniform manner. Copies of this blank may be obtained from A. B. Moehlman, Director of Reference and Statistics, Board of Education, Detroit.

Part I is the usual form employed by the Federal Bureau of Education. The organizations which proposed Part II have been concerned merely with the financial section of the Federal Report. In other words, Part I stands as it always has.

We have recently received from Superintendent C. B. Smith of Princeton, Illinois, a copy of a test in use in that county, and known as the Bureau County Standard Test in Speech Error Correction. It was devised by Superintendent Smith at the request of the County Superintendent in order to check the extent to which the teachers who cannot have much supervision have successfully taught the course of study. The test consists of 20 sentences for each grade, a number of which contain an incorrect language form. The child in taking the test is to place an *r* after each correct form and an *x* after each incorrect form. The score is the number of right minus the number of wrong answers multiplied by 5.

This is an interesting piece of problem work. The test is not completely diagnostic since children may be able to recognize some incorrect forms without being able to correct them. This, however, is not a serious fault considering the purpose for which the test was devised.

It is a most hopeful sign when local school officials recognize the value as instruments of supervision of more or less uniform tests which lend themselves to objective scoring. The efforts of such men as Superintendent Smith are certainly to be encouraged, and the assistance of those who have training and experience in the work of devising educational tests will most gladly be given.

Dr. H. L. Camp, Assistant High School Visitor for the University of Illinois presented some data on the use of general intelligence tests before the National Association of High School Supervisors and Instructors at the Chicago meeting of the N. E. A. which is worth a wider distribution than the membership of that Association.

Dr. Camp endeavored to ascertain the extent to which general intelligence tests were being used throughout the country. He received reports from 80 percent of the states; and while the men addressed expressed themselves enthusiastically in favor of tests, their reports indicate a rather surprising lack of systematic work. No state in the Union has conducted anything like a state wide survey of intelligence. The nearest approach to it was in Indiana where the Indiana Intelligence Scale, Schedule D was given to 6,188 seniors in 320 high schools.

The Terman Group Intelligence Test was mentioned as having been used by workers in 15 states; the Otis in 14; The National in 13; and the Army Alpha in 10. Less frequent references were made to the Illinois Examination; Chicago Examination; Indiana Intelligence; Thorndike High School Examination and College Entrance Examination; Army Examination and Tests by Haggerty, Miller, Trabue, and Myer.

A "deficiency survey" of the high-school pupils who have been reported as doing unsatisfactory work during the first semester of this school year has been made by Principal H. C. Koch of Mt. Vernon, Ohio. Little agreement was found between pupils and teachers as to the reasons for low marks. A larger percent of deficiency was reported for pupils from the neighboring rural school than for those from the local schools.

The intelligence score of the 256 unsatisfactory pupils was compared with the scores reported for 5520 high-school pupils in the Manual of the Army Alpha Intelligence Scale. Only one-quarter of these unsatisfactory pupils had intelligence scores equal to or above the median of this larger group. No cases in the army group were rated D or D—, while one percent of the unsatisfactory pupils at Mount Vernon were so rated.

This study disclosed several needs among which were: (1) classification of students on the basis of mentality; (2) gradation of requirements within courses from mere fundamentals to a high degree of enrichment; (3) a survey of pupil's needs, habits, and home conditions; (4) a thorough campaign of education of parents in home-school responsibility.

The Tax Campaign in St. Louis

March 28, 1922, saw the close of a campaign for the maintenance of the tax rate for schools of 85¢ on the $100.00 in St. Louis that in many respects was notable.

On that day the tax payers of the city of St. Louis by a vote of little less than three to one approved the maintenance of the existing tax rate of 85¢ on the $100 which actually was a continuance of an increase of 25¢ over the rate of 60¢ allowed by the State Constitution. In order to maintain that rate, public-spirited citizens combined with public school teachers and pupils to wage what proved to be a resistless campaign of publicity and persuasion. The voter was appealed to from every angle. A Committee on Statistics and Research which had been organized early in the campaign compiled quantities of data upon school questions of almost every conceivable variety. These data were freely drawn upon by other committees and by speakers with the purpose of presenting the claims of the public schools to the men and women of St. Louis in the best possible light.

Another committee which performed a remarkable piece of work was that on posters and display. Prompted by this committee boys and girls in the high schools and pupils in the elementary schools from the first to the eighth grade prepared striking posters in colors with colored crayons and crayola. These posters were distributed over the city and were displayed from the windows of private homes as well as of department stores, factories, and wholesale business houses.

The committee on speakers completely covered the city. It arranged for meetings in every public school. It sent five-minute speakers to factories and other industrial plants during the noon hour. It secured the approval of the Ministerial Alliance to an educational Sunday, when the churches appealed to their membership for support of the school tax rate. It secured the approval of countless organizations, both those identified in some way with the school and its life or unrelated organizations, like the City Central Committees of the various political parties, labor unions, civic and philanthropic and miscellaneous business clubs and societies.

The Committee on Publicity after securing material from the Committee on Statistics and Research worked it up and, with the assistance of the president of a prominent advertising company, phrased its appeal in a form approved by the best advertising practice. Through this committee the schools were circularized, a campaign of publicity was pursued in the newspapers by means of news articles and advertisements and with the cooperation of moving picture houses patrons of the screen were for four weeks nightly greeted with appeals based on civic pride and love of the public school to "Vote YES and Scratch NO, March 28th, 1922."

One of the most telling features of the campaign was the composition and singing of school songs advocating the approval of the tax rate. This succeeded in bringing to the surface an astonishing amount of latent talent which otherwise the schools would not have known they possessed.

Comment on Professor Knight's Recent Article

May I be permitted a word of comment on the conclusions reached on the basis of data presented in an article in your Journal for March, 1922, by F. B. Knight, entitled: "Qualities Related to Success in Elementary School Teaching."

I am particularly interested in the explanations offered by Mr. Knight for the finding of zero correlations between two sets of "ascertainable facts concerning the teachers" and "the degree of success they were obtaining in actual class-room work" namely "(1) professional study during service, and (2) mental tests."

I. I do not think that Mr. Knight should be "confused by the absence of relationship" between the "amount of summer school and Saturday work in educational courses" and "skill in teaching." The expected correlation would probably be found to exist if the persons who took this work did so out of motives of intellectual curiosity and professional ambition originating within themselves. The fact that the correlation does not exist, seems to me to be due to a situation easily observable at any summer session, namely: that an important proportion of teachers attending these sessions do so under compulsion of various kinds not related to the motives above referred to.

II. It seems remarkable that Mr. Knight should offer in explanation of the absence of correlation between the mental test and teaching ability, the opinion that "too much intelligence might also be a handicap." A too active mind, "he thinks, might well be so constantly aware of unsolved problems . . . . that superior intelligence would get in its own way in the rather humdrum and steady grind of elementary-school teaching."

It seems difficult to believe that superior intelligence, which in all likelihood is also creative and contributing intelligence, should run away from unsolved problems. Is it not the business and tendency of such intelligence to assist in the solution of such problems? May one be permitted to refer to the author's attempt to solve one of the obviously difficult problems of educational administration as an apt example of the contrary idea?

It would seem to the writer that the reason for the absence of correlation between the mental test and teaching ability is to be found in the fact that the higher up we go in the scale of intelligence, the less reliable does the quantitative score become, and the less important do the apparent distinctions in quantitative differentiation become. To give an illustration: in the Stanford-Binet scale, a distinction of fifteen points does not always have the same significance. Between I. Q. 65 and I. Q. 80 it means by almost any definition, the difference between feeblemindedness and normality. On the other hand, between I. Q. 145 and 160, the difference is of comparatively negligible significance. As between persons of that intelligence status, the merely quantitative measure has become useless. To quote another example: A group of children all of them having I. Q.'s between 125 and 140, is from the point of view of the available quantitative means of distinguishing between them, a homogeneous group, but they differ qualitatively. This becomes very obvious when one attempts to apply the Stanford-Binet test to the measurement of very superior adults. One might find ten adults all of whom answer every single question of the Stanford-Binet scale satisfactorily. In that case the scale simply does not function; it does not measure; it does not differentiate. Furthermore, an attempt to extend it would probably prove futile.

The demand of the moment is for an intelligence measure which will differentiate qualitatively between members of a quantitatively homogeneous group. This is the next great problem of applied psychology.

When one bears "in mind . . . . that all the teachers involved in this study were in the *upper portion of the total range of intelligence,*" it is *not at all* "evident that a mental test is not a good selective device as far as elementary school teaching is concerned." The thing that is evident is that when the distribution of intelligence falls within a comparatively small range, and that range the upper portion of the total distribution, it fails to function as a method of differentiation between members of the group. A qualitative differentiation tending to discover within a quantitatively homogeneous group those members whose intelligence tends in the direction of good teaching would function. Is it not possible that the Knight-Franzen Test for elementary school teachers showed some correlation because it does in fact tend toward being such a qualitative intelligence test?

John Lewis Horn

Mills College

A Prophet in His Own Country

Mr. J. L. Stenquist, whom our readers will doubtless remember as the author of two recent articles in the Journal, articles in which he stresses the unreliability of individual intelligence scores and makes out a case for the child of low I. Q., has not been saved from attack by his failure to join the ranks of those who uncritically accept the results of intelligence testing. Principal Gilbert J. Raynor of the Commercial High School of New York falls foul of Mr. Stenquist in the April 1922 number of the "Bulletin of High Spots." In his attack Principal Raynor succeeds in exhibiting a surprising degree of intellectual discontinuity. Mr. Stenquist, at the suggestion of Superintendent Tildsley, (who, according to our last information, was a district superintendent assigned to high schools) arranged to give a series of intelligence tests at the Annex of Commercial High School to the boys who were failing in all subjects. At the outset Mr. Raynor quotes Mr. Stenquist's report as follows:

"These boys, as a group, are of exactly average intelligence."

"It is clearly shown that several boys have marked ability, not only in mechanical ability, but in intelligence . . . ," "As compared with men in the army who are clerks, this whole group ranks much higher."

"Other causes are apparently responsible for the failure of these boys in the first term of high school."

After the above quotations Mr. Raynor declares that these conclusions "are so wide of the facts as to be almost humorous." From this point on the argument seeks to show first, that the tests were inadequate. In suggesting a more extensive program, Mr. Raynor not only suggests the use of a "comprehensive series such as that given in Whipple's Manual" but also suggests an investigation of family history, medical history, a physician's examination and a dip into the pupil's previous school record. This is the first *non sequitur*. We wonder whether the writer would have advised Mr. Stenquist to give *all* the tests which are described in the two volumes of Whipple Manual. Either Mr. Raynor doesn't know the scope of Whipple's Manual or he is suggesting a testing program of such magnitude in order to discredit any testing program. But this is not the point in the non-sequential character of Mr. Raynor's thinking. He wants family and medical history, and a physician's examination. This is not what Mr. Stenquist was asked to do. According to Mr. Raynor's own words it was at the suggestion of Dr. Tildsley that Mr. Stenquist "arranged to give a series of intelligence tests." If Mr. Raynor wanted a physician's examination or if he wanted an investigation of family history or of each pupil's previous school record, these were other matters and for the most part the school records should have contained the information. To obtain such data was certainly no part of the task which Mr. Stenquist undertook—namely the giving of a series of intelligence tests.

A second defect in Mr. Raynor's line of argument gathers about his misquoting from Mr. Stenquist's report. As above stated the report reads "the boys, as a group, are of exactly average intelligence." At this point, however, Mr. Raynor makes Mr. Stenquist's report read "these boys are of exactly average intelligence." The omission of the words "as a group" is well suited to the development of the line of attack which follows. Presumably Mr. Stenquist had shown that the average or median intelligence quotient for the group was in the neighborhood of 100. This would entirely justify the statement as made in his report but it would not justify the garbled quotation which Mr. Raynor used from the report.

But the more astonishing feature of Mr. Raynor's paper is that it is a precise justification of Mr. Stenquist's conclusion that "other causes are apparently responsible for the failure of these boys in the first term of High School." Mr. Raynor relates how he interviewed personally each one of the boys whom Mr. Stenquist had tested. This was undoubtedly a good thing to do and we should have supposed he might have done it as part of his job without the pugnacious incentive which apparently actuated him. The information he obtained from these interviews and from interviews with the parents was most illuminating. He finds that any suggestion that these failures were due to poor teaching is ridiculous. So far as he has quoted Mr. Stenquist there is no suggestion to this effect. The following facts were disclosed concerning some of these boys: (1) an abnormally large amount of absence, (2) a disinclination to attend high school, (3) unwillingness to study, (4) "pitiful" home conditions, (5) poor physical conditions, (6) incorrigibility, and (7) physical exhaustion due to too long hours of outside work.

In piling up these "other causes" Mr. Raynor is unconsciously corroborating Mr. Stenquist's view that these other causes were "apparently responsible for the failure of the boys." Indeed it would have been easy for Mr. Raynor, using exactly the same data, to have written an article in support of Mr. Stenquist rather than in opposition to him. In truth it is the animus of this paper rather than the facts adduced which justify us in calling it an attack. Undoubtedly there is something behind this. Perhaps Dr. Tildsley's suggestion that intelligence tests be given was not acceptable to the principal. So far as the suggestion implied that a complete or even satisfactory determination of the causes of failure on the part of these pupils would be thus afforded we should agree with the principle entirely. So also would Mr. Stenquist who has been one of the foremost among the makers and users of tests to point out their limitations. But Mr. Stenquist is in the employ of the New York Board of Education. If a superintendent suggests that he give a series of intelligence tests, all that can be expected either by the superintendent or anyone else is that the intelligence tests be well selected, carefully given and rated, and sanely interpreted. If an expert not in the employ of the Board of Education had been brought in to do this testing, it is probable that his work would have been more highly respected. A prophet is not without honor save in his own country.

Denton L. Geyer of the Chicago Normal School has been using the 21st Yearbook of the National Society for the Study of Education as a textbook in a course on intelligence testing. He has conceived the idea of the development of a standardized test on the material of this book, to be used as a final examination for the course. He has devised such a test and on the basis of his results during the past quarter he secured a correlation of 0.30 with the average scholarship of his students for the previous semester.

Since a great many teachers in universities, colleges, and normal schools will doubtless use this book to a great extent in the courses this summer, they may be interested in communicating with Mr. Geyer relative to the test which he has devised.

Superintendent F. H. Bair of Colorado Springs reports the following salary schedule for the current year: minimum salary $1,200; maximum for teachers with two years preparation beyond high school, $1,950; for those with three years beyond high school, $2,140; for those with four years but without a degree, $2,400; for college graduates, $2,700; and for those with the M. A. Degree, $2,800. Salaries are the same in all grades for the same degree of preparation and teaching experience. To encourage professional growth, the Board provides a $100.00 annual increment to the salary for attendance at a standard summer school. As a result of this provision 50% of the teachers attended last summer and more will attend this year.

The following letter was lately received by an applicant for the superintendency of the schools in a community of 7,000. It is certainly surprising that there are communities of that size in the United States in which the attitude shown in this letter is taken.

Mr. ——————
New York City.
Dear ——————:

Your application for Superintendent of the ———— schools has been received, together with your references. It is so uncertain as to whether we will have a school that I deem it but fair to you to write you about it. Unless some radical change takes place in the disposition of our people, it looks very much as if we will have no school during the ensuing year. I am filing your application and references for future consideration. I am frank to say, however, that it will be very unwise to depend on securing a position in our schools for the coming year, owing to the circumstances related.

Yours very truly,
——————————, Secretary.

Salaries and Cost of Living

Mr. Don C. Rogers of the Bureau of Educational Service, State University of Iowa, has recently given out the results of a study on salaries and cost of living affecting superintendents, principals, and teachers in that state. Salaries for 1921-22 showed increases over 1920-21 in most types of work and groups of cities. When both sexes were compared and cities of all sizes considered together, median salaries for the past year showed as follows: superintendents, an increase of $123.00; high-school principals, an increase of $12.00; high-school teachers, an increase of $36.00; grade teachers, a decrease of $5.00.

The average cost of board for teachers in April, 1922, (on the basis of returns from 80 cities and towns in the state) was $7.04 per week as compared with $7.32 for April, 1921. On the other hand, the average payment for rent was $2.78 as against $2.75 the previous year. This indicates that so far as these two items of cost of living were concerned, the H. C. L. for an average teacher decreased $1.00 per month.

In view of the insistent demand in many quarters for reduction of teachers' salaries for the coming year on the basis of decreased cost of living, this study is pertinent. When these figures are considered in connection with the fact that teachers' salaries did not begin to rise until the school year 1918-1919, one cannot logically agree that reductions in salaries are justified.

Coordination of Research

A sub-committee of the N. E. A. Commission on the Coordination of Research Agencies met in Chicago on April 29 and 30, 1922. The following were present: Newlon (chairman), Bobbit, Briggs, Buckingham, Cary, Courtis, Dickson, Norton, Pickell, Studebaker, and Theisen.

The decisions of the meeting are important to those interested in educational research. It became at once evident that the creation at the N. E. A.

headquarters of a research department had given the Commission an opportunity to be definite in its recommendations. The agency by which the methods approved by the committee were to be put into effect seemed naturally to be this new research department. One might reasonably feel a bit sorry for Director Norton. To be called upon to "function" in so many different ways so soon after being appointed to a newly-created position seems rather rough treatment. Mr. Norton, however, appears to be determined to meet the wishes of the committee as far as it is humanly possible. If he fails to put into effect the many suggestions of the committee, it will probably be because of a lack of resources.

The sense of the meeting was expressed in a number of particulars, to a few of the most important of which we refer.

It was suggested that a working reference library should be set up in the research department and that it should be catalogued and indexed for use in furnishing data to school people.

It was further suggested that the department act as a clearing house, and that as such it should gather information, formulate statements of current and future problems, and render service to school people. This service, it was suggested, might be largely on a self-supporting basis. Committees were appointed, subject to ratification, as follows: (1) to work out a classification of educational activities and functions (Briggs, Courtis, and Buckingham); (2) to ascertain the present and probable research activities of departments of the N. E. A. and to report a plan for coordinating the work of these agencies (Dickson, Studebaker, and Melcher). Committees were also recommended as follows: (1) an advisory committee to work in cooperation with the research department in formulating policies and plans of the department; (2) a committee on publicity to create through every kind of publicity medium a correct impression of present-day educational ideals, methods, and tendencies. It was added that the work of the publicity committee should be adequately financed.

Much discussion of the troublesome topic of questionnaires was held, especially as to their abuse and the duplication involved in them. Feeling that the N. E. A. through its research department could do something to control the situation, the members of the committee passed a resolution making the research department responsible for the coordination of the information-collecting activities of educational agencies by formulating, issuing, tabulating, and making available through publication or by other means, standard questionnaires on topics of major importance for use by those interested in educational problems. In further definition of this position, the following procedure was recommended:

(1) The Director of the Research Department about May 15th should send communications to a selected group of varied educational agencies requesting a statement of the major problems in which they are interested.

(2) Based upon the replies to this communication, the Director should prepare a list of the problems of most common interest and submit the list not later than June 15th to the members of the Commission

on the Coordination of Research Agencies for criticisms and suggestions.

(3) Not later than August 15th the Director should submit to the members of the Commission for their criticism and suggestions detailed questionnaires covering the topics selected by him together with specific plans for conducting investigations based upon these questionnaires either through the central research department or by allocation to other responsible persons and agencies.

(4) The Director should issue in the September or October number of the Journal of the N. E. A. (in the name of the Commission if advisable) a statement of the general plan together with a calendar showing the approximate dates upon which the results of the various investigations may be secured from the central research office or from the cooperating agencies.

(5) As rapidly as proves feasible this plan for coordinating the collection, compilation, and distribution of desired information should be extended to include a wider and wider range of problems.

At the conclusion of its deliberations the committee passed the following resolution: "The Commission, realizing the broad scope and large expense connected with any adequate plan of organizing the investigational activities of the nation, recognizes that no agency for the coordination of such activities will be adequate except that which has the prestige of a Federal Department of Education as proposed in the Towner-Sterling Bill."

Time Distribution in Some Louisiana Schools

The Callender Laboratory of Newcomb College studied the time distribution of about 25 percent of the city school systems and 20 percent of the rural school systems of Louisiana. The purposes of the study were: to obtain a basis for a study of the most effective time distribution; to learn what subjects were allotted the greatest or least amount of time and emphasis and hence, upon which most or least time and energy were expended; and to find the relation between the amount of time and the results of instruction as revealed by standardized tests. In studying the problem, the definition of subjects followed the scheme used by Holmes in the Fourteenth Yearbook of the National Society for the Study of Education.

In comparison with the findings of Holmes, the Louisiana schools give less total time to each subject than the northern cities with the exception of opening exercises, penmanship, and physical training. The rural schools give the same amount of time to penmanship as northern cities and more to sciences. The Louisiana city schools give a greater proportion of the entire school time to recitations than do northern cities except in history and sciences. The rural schools devote a greater percent of time to science and history than do the schools reported by Holmes. The variability of time distribution is less in every subject than in the schools of the Holmes' study.

All schools offer spelling, reading, and arithmetic through all the grades and over 90 percent of the rural schools also give language, geography, and history. All city schools give language, and 90 percent give history, but only 60 percent report geography. Penmanship is reported by 85 percent of the rural schools and by all city schools. It is continued through the grades by the majority of the rural schools, but beyond the fifth grade by only 30 percent of the city schools. The variability in the amount of time allotted to each subject in the different grades seems ample reason for raising the question as to what constitutes the basis on which time distribution was made.

When the time allotted is considered according to subjects, two features stand out prominently: (1) variability, and (2) the regular appropriation of the larger portion of the time by the formal subjects. City schools show least variability and smaller rural schools the greatest. Country children spend 92 hours more per year in school than do the city children. Recitation periods are longer in city schools and study periods are longer in the rural schools. If the average time for all grades is taken as a basis, a pupil passing through the rural schools will spend the greatest part of his school career on reading with the following activities in decreasing amounts: recess, arithmetic, language, science, penmanship, geography, history, spelling, miscellaneous, community work, agriculture, opening exercises, industrial training, and drawing. Similar arrangement for a city child would be: reading, arithmetic, language, physical training, penmanship, miscellaneous, spelling, recess, geography, opening exercises, drawing, history, community work, and agriculture.

The time allotted for study periods shows a similar concentration; for 65 percent of the study time in the rural schools is given to reading, language, spelling, penmanship, and arithmetic; while in the city schools 85 percent of the study time is given to reading, language, spelling, and arithmetic. The city schools assign more of both recitation and study time to the formal subjects than do the rural schools.

COMPARISON OF LARGEST AND SMALLEST TIME ALLOTMENT

Subjects	Recitation		Study	
	Rural	City	Rural	City
Reading	25	7	42	10
Language	50	4	19	2½
Spelling	30	8	24	2
Penmanship	8	2	12	
Arithmetic	36	4	18	10
Geography	8	3	12	2
History	7	7	15	2
Science	20	2	12	1½
Drawing	18	10		
Agriculture	4			
Community Work	9	1¼	14	1¼

The data are read thus: in reading, for example, the rural school giving the most time to this subject allows 25-times as much time to it as the school giving the least amount of time; the city school making the largest allotment devotes 7-times as much time as the one making the least time allotment. The largest amount of study time for reading in rural schools is 42-times the smallest, and in city schools, the largest is 10-times the smallest.

Where such variability exists in the expenditure of time and energy, great differences should also exist in the achievement. Only through the use of standardized tests are we able to determine the extent to which variability in achievement justifies this great variability in time expenditure. May it be possible in the near future to plan a more economical and effective time-distribution based on such test results!

DAGNY SUNNE

Newcomb College

The Program at San Jose

A splendid program was rendered on May 19 and 20 at the Conference on Educational Research and Guidance of the San Jose, California, State Teachers College. A number of the best research people on the Pacific Coast were on the program. President Kemp made the opening address on the morning of May 19 and Dr. Proctor talked on "Vocational Guidance." Dr. Franzen of Des Moines seems to have been the only person from east of the Rockies who contributed to the program.

A remarkable series of papers was presented in the afternoon on the general topic of gifted children. Professor Terman led the discussion presenting a paper on "The California Survey of Gifted Children." Dr. Dickson discussed "The Treatment of Gifted Children in Oakland." Dr. Dorothy Yates gave her "Impressions from Two Years Spent in Testing Gifted Children." Dr. Grace Fernald presented material on "Teaching Results in the Case of Bright Children"; and James C. DeVoss gave a talk along somewhat the same lines entitled "School Success of Gifted Children." The discussions ended with a paper by Dr. Esther Gaw on "Testing for Musical Talent."

Dr. J. Harold Williams delivered an evening address on "Some Aspects of Delinquency."

On May 20 a conference on school surveys was held, the discussion being led by Dr. Sears. He addressed the conference on "Fundamentals in Surveying" and was followed by Dr. Hart on "Building Surveys and Building Programs," Dr. Sutherland on "The Teacher and School Surveys," and Dr. Freeland on "The Principal and the School Survey." Dr. French of the Bureau of Research at the University of California and Mr. DeVoss of the Bureau at San Jose State Teachers College reported concerning their organizations.

The final conference was on "Adjustment." Dr. Dickson led the discussion, his formal topic being "Mental Tests and Their Significance in High School." At this session Dr. Breitwieser reported on the "Mental Survey of Modoc County," and Dr. Williams on "Mental Tests and Delinquency." State Superintendent Will C. Wood spoke on "Adjustments in California."

This was evidently worth listening to. The folks on the Pacific desire to organize a division of the Educational Research Association. We feel that they should be encouraged in this plan. They have the strength both in numbers and in quality to make a strong organization, and their isolation prevents their regular attendance at the annual meetings of the national organization.

Predictive Value of Scholastic Marks

By the help of people who know the graduates of the Kansas State Agricultural College three lists of 50 graduates each were selected for study. List one was made up of those who are highly successful, each in his own vocation; list two of those who have made an average success; and list three of those who are ranked below average in success. All these graduates have been out of college ten or more years.

The scholastic marks given to these 150 graduates during their four years in college were tabulated with the following results, E representing the highest and F the lowest marks:

	E's	G's	M's	P's	F's
	% of total marks	% of total marks	% of total marks	% of total marks	% of total marks
Highly Successful	29	46	19	6	0
Average	4	14	52	23	7
Below Average	1	5	19	62	13

Tentative Conclusions:

1. If 75% or more of the scholastic marks of a graduate of the Kansas State Agricultural College are E's and G's he will be likely to be highly successful in his life work.
2. If 75% or more of his scholastic marks are M's and P's he will be likely to make an average success in his life work.
3. If 75% or more of his scholastic marks are P's and F's he will probably rank below the average in his life work.

EDWIN L. HOLTON

Kansas State Agricultural College

In a pamphlet published by the Children's Bureau of U. S. Department of Labor Dr. Wm. Healey reports a study of delinquency among children.

"A table of 1,200 juvenile offenders distributed according to their 'intelligence quotient' shows that the great bulk of the repeated offenders come from children of 80-100 I. Q. There are as many offenders between 110 and 120 as there are between 60 and 70 I. Q. There are more offenders between 100 and 110 than there are between 70 and 80 I. Q.

"It is conceded that, in proportion to their number in the community, mentally defective children furnish perhaps ten times their quota of incipient criminals, but in the aggregate the dull-normal group actually furnishes the bulk of children with criminal tendencies.

"Dr. Healy finds that mental capacity or general intelligence is only one of ten factors that may influence a child toward unsocial behavior. The one greatest factor he appears to believe is that of the child's 'mental content, idealation, or imagery'."

Educational Research Association

(FRANK N. FREEMAN, *Secretary and Editor*)

The space assigned to the Educational Research Association is taken up this month with the membership list. Two or three news items which have come in will be put in the next issue. The secretary takes this opportunity again to ask for communications. Three types are suggested: first, personal items of news; second, brief reports of work in progress or of completed studies which are not extensive enough for an article; third, critical discussions of most questions. Let us have a little real criticism. If anybody cannot stand professional, objective criticism he is not a scientist.

It is appropriate in connection with the publication of the membership list to repeat the provision concerning the qualifications for membership. The constitution provides for election by the executive committee "on the presentation of satisfactory evidence in the form of published or unpublished studies which show ability to arrange, organize, and to conduct research . . . " The executive committee usually regards publication as a requisite. If possible, four copies of publications should be submitted with applications. Blanks for application have been printed and will be sent to any member desiring to nominate somebody else for membership or to any person desiring to apply.

As the membership roll lengthens it becomes more difficult to keep it up to date. The secretary especially invites corrections of the following list.

Since there are occasional requests for copies of the membership roll a few reprints of the present list are being made. They may be obtained by addressing the secretary upon the payment of fifteen cents.

LIST OF MEMBERS OF THE EDUCATIONAL RESEARCH ASSOCIATION

HONORARY MEMBERS

Ayres, L. P., Vice-President Cleveland Trust Co., Cleveland, Ohio.

Cattell, J. M., Garrison-on-Hudson, N. Y.

Chadsey, C. E., Dean of School of Education, University of Illinois, Urbana, Illinois.

Coffman, L. D., President University of Minnesota, Minneapolis, Minn.

Hall, G. Stanley, Clark University, Worchester, Mass.

Hanus, P. H., Professor of Education, Harvard University, Cambridge, Massachusetts.

Judd, C. H., Director of School of Education, University of Chicago, Chicago, Illinois.

Rice, J. M., Germantown, Pennsylvania.

Russell, James E., Teachers College, Columbia University, New York City.

Russell, W. F., Dean of School of Education, University of Iowa, Iowa City, Iowa.

Thorndike, E. L., Professor of Educational Psychology, Columbia University, New York City.

Wissler, E. C., Washington, D. C.

REGULAR MEMBERS

Alexander, Carter, 525 West 120th St., New York City.

Anderson, H. W., Assistant Director, Educational Research, 50 Broadway, Detroit, Michigan.

Ashbaugh, E. J., Assistant Director, Bureau of Educational Research, Ohio State University, Columbus, Ohio.

Baldwin, B. T., Director, Child Welfare Research Station, University of Iowa, Iowa City, Ia.

Ballou, F. W., Supt. of Schools, 1340 Girard St., Washington, D. C.

Barthelmess, Harriet M., Teachers College, Columbia University, 106 Morningside Drive, New York City.

Beeson, M. J., University of Colorado Cooperative Extension, Grand Junction, Colorado.

Brainerd, Mrs. Margaret, Director of Educational Research, Martins Ferry, Ohio.

Branson, E. P., Director, Bureau of Educational Research, City Public Schools, Long Beach, California.

Breed, F. S., University of Chicago, Chicago, Illinois.

Bright, Ira J., Supt. of Schools, Leavenworth, Kansas.

Brown, Emma M., Director, Bureau of Research, Public Schools, Denver, Colorado.

Brueckner, L. J., City Teachers College, Detroit, Michigan.

Buckingham, B. R., Director, Bureau of Educational Research, Ohio State University, Columbus, Ohio.

Buckner, C. A., University of Pittsburgh, Pittsburgh, Pa.

Burgess, Mrs. May A., Dept. of Education, Russell Sage Foundation, 500 West 122nd St., New York City.

Buswell, Guy T., University of Chicago, Chicago, Illinois.

Caldwell, O. W., Lincoln School, Teachers College, Columbia University, New York City.

Charters, W. W., Carnegie Institute of Technology, Pittsburgh, Pa.

Clark, J. R., Teachers College, Columbia University, New York City.

Cobb, Margaret, Institute of Educational Research, Teachers College, Columbia University, New York City.

Colvin, S. S., Brown University, Providence, Rhode Island.

Counts, George S., Yale University, New Haven, Connecticut.

Courtis, S. A., Dean, Teachers College, 246 Eliot St., Detroit, Michigan.

Coxe, W. W., Graduate Student, Ohio State University, Columbus, Ohio.

Cram, F. D., Iowa State Teachers College, Cedar Falls, Iowa.

Cubberley, E. P., Leland Stanford University, Stanford University, Calif.

Daley, H. C., Director, Survey Department, Highland Park, Michigan.

Dalman, M. A., Director, Educational Research, Board of Education, Indianapolis, Indiana.

Davis, Helen, Director, Educational and Mental Measurements, Jackson, Michigan.

Dawson, C. D., Assistant Supt. of Schools, Grand Rapids, Michigan.

DeVoss, J. C., Director, Bureau of Research and Extension, State Teachers College, San Jose, California.

Dickson, V. E., Director, Bureau of Research and Guidance, Oakland, Cal.

Elliott, Charles, Rutgers College, New Brunswick, New Jersey.

Everly, L. L., Ass't Supt. of Schools, St. Paul, Minnesota.

Flemming, Cecile W., Teachers College, Columbia University, 519 W. 123rd Street, New York City.

Fordyce, Chas., Director, Bureau of Educational Measurements, Teachers College, Lincoln, Nebraska.

Franzen, R., Director of Research, City Public Schools, Des Moines, Ia.

Frazier, Geo. W., Director, Bureau of Educational Research, Public Schools, Denver, Colorado.

Freeman, Frank N., University of Chicago, Chicago, Illinois.

Gates, Arthur I., Teachers College, Columbia University, New York City.

Gray, C. T., University of Texas, Austin, Texas.

Gray, W. S., University of Chicago, Chicago, Illinois.

Greene, H. A., Bureau of Educational Service, Extension Division, University of Iowa, Iowa City, Ia.

Guy, J. F., Director, Research and Measurement, 720 Fulton Bldg., Pittsburgh, Pennsylvania.

Haggerty, M. E., Dean, College of Education, University of Minnesota, Minneapolis, Minnesota.

Harlan, C. L., Department of Educational Measurements, Lewiston State Normal, Lewiston, Idaho.

Heck, A. O., Director, Bureau of Educational Research, Public Schools, Akron, Ohio.

Henmon, V. A. C., Dean, School of Education, University of Wisconsin, Madison, Wisconsin.

Henry, Mary B., Director, Research and Guidance, Santa Anna, Calif.

Herring, J. P., Director, Bureau of Educational Research, State Normal School, Bloomsburg, Pennsylvania.

Hollingworth, Mrs. Leta, Teachers College, Columbia University, New York City.

Horn, E., Director, University Elementary School, State University of Iowa, Iowa City, Iowa.

Hughes, W. H., Supervising Principal, Claremont, California.

Inglis, Alexander, Harvard Graduate School of Education, Cambridge, Massachusetts.

Johnson, O. J., Asst. Director of Research, St. Paul Public Schools, St. Paul, Minnesota.

Kallom, Arthur, Asst. Director, Bureau of Researth, Huntington Avenue, Boston, Massachusetts.

Kelley, T. L., Leland Stanford University, Stanford University, Cal.

Kelly, F. J., Dean, Administrative Research, University of Kansas, Lawrence, Kansas.

King, L. A., University of Pennsylvania, Philadelphia, Pa.

Kirby, T. J., University of Iowa, Iowa City, Ia.

Knight, F. B., University of Iowa, Iowa City, Ia.

Koos, L. V., University of Minnesota, Minneapolis, Minnesota.

Kuhlmann, Frederick, Director, Research Bureau, State Board of Education, 1079 Dayton Ave., St. Paul, Minnesota.

McCall, Wm. A., Teachers College, Columbia University, New York.

McCallie, J. M., Director, Bureau of Educational Research and Efficiency, Trenton, New Jersey.

McCory, J. L., Asst. Supt. of Schools, Omaha, Nebraska.

McMillan, E. E., Central High School, Omaha, Nebraska.

Mallory, Clara, Director, Educational Research, 1016 Liberty St., Beaumont, Texas.

Manuel, H. T., Colorado State Normal, Gunnison, Colorado.

Melcher, George, Director, Bureau of Educational Research, Library Bldg., Kansas City, Missouri.

Miller, W. S., University of Minnesota, Minneapolis, Minnesota.

Moehlman, A. B., Director, Div. of Statistics, City Schools, Detroit, Mich.

Monroe, W. S., Director, Bureau of Educational Research, University of Illinois, Urbana, Illinois.

Morrison, J. C., Specialist in Educational Measurements, State University, Albany, New York.

Morton, R. L., Ohio University, Athens, Ohio.

Myers, Anna G., Library Bldg., Kansas City, Missouri.

Murdock, Katherine, Polk, Pa.

Myers, Gary C., Cleveland School of Education, Cleveland, Ohio.

Neale, M. A., University of Minnesota, Minneapolis, Minnesota.

Nifnecker, Eugene A., Director, Bureau of Research and Statistics, 500 Park Ave., New York City.

Norton, J. K., Director, Research Dept., N. E. A., 1201 16th Street, N. W., Washington, D. C.

O'Brien, F. P., Director, Bureau of Service, University of Kansas, Lawrence, Kansas.

O'Hern, J. P., Ass't Supt. of Schools, Municipal Bldg., Rochester, N. Y.

Osburn, W. J., Director, Educational Measurements, State Dept. of Public Instructions, Madison Wisconsin.

Otis, A. S., World Book Co., Yonkers, New York.

Packer, P. A., State University of Iowa, Iowa City, Iowa.

Peterson, Joseph, Peabody College for Teachers, Nashville, Tenn.

Phelan, W. W., University of Oklahoma, Norman, Okla.

Pressey, Mrs. L. C., Ohio State University, Columbus, Ohio.

Pressey, S. L., Ohio State University, Columbus, Ohio.

Proctor, Wm. M., Leland Stanford University, Stanford Univ., Cal.

Race, Henrietta V., Director, Bureau of Educational and Mental Measurements, Board of Education, Youngstown, Ohio.

Rice, D. Edgar, Apt. 22, The Kenesaw, Washington, D. C.

Richardson, J. W., Director, Educational Research, Public Schools, Hibbing, Minnesota.

Rinsland, H. D., Director, Bureau of Research, City Schools, Ardmore, Okla.

Rugg, H. O., 646 Park Ave., New York City.

Sexton, E. D., Asst. Supt., Department of Reference and Research, Newark, New Jersey.

Smith, L. O., Asst. Supt., 508 City Hall, Omaha, Nebraska.

Staker, M. P., Director, Bureau of Educational Measurements, Aberdeen, South Dakota.

Stenquist, J. L., Bureau of Reference Research and Statistics, Public Schools, New York City.

Stevenson, P. R., Research Associate, Bureau of Educational Research, Ohio State University, Columbus, Ohio.

Strayer, G. D., Teachers College, Columbia University, New York City.

Sutherland, A. H., Director, Dept. of Psychology and Educational Research, Los Angeles, California.

Sutton, C. W., South High School, Cleveland, Ohio.

Swift, F. H., College of Education, University of Minnesota, Minneapolis, Minnesota.

Talbert, W. E., 455 Flood Bldg., San Francisco, California.

Terman, L. W., Leland Stanford University, Stanford Univ., Cal.

Theisen, W. W., Director, Bureau of Educational Research, Public Schools, Cleveland, Ohio.

Torgerson, T. L., Director, Educational Research, City Schools, West Allis, Wisconsin.

Toops, H. A., Teachers College, Columbia University, New York City.

Trabue, M. R., Teachers College, Columbia Univ., New York City.

Updegraff, H., University of Pennsylvania, Philadelphia, Pa.

Van Wagenen, M. J., College of Education, University of Minnesota, Minneapolis, Minnesota.

Washburne, C. W., Supt. of Schools, Winnetka, Illinois.

Welles, J. B., Supervising Principal, Roslyn Heights, Long Island, N. Y.

Whipple, G. M., University of Michigan, Ann Arbor, Michigan.

Wight, A. G., Asst. Supt. in charge of Measurements, Chicago Public Schools, Chicago, Illinois.

Williams, L. A., University of North Carolina, Chapel Hill, N. C.

Williams, J. Harold, Bureau of Juvenile Research, Whittier, Cal.

Williams, Osburne, U. S. Public Health Service Hospital No. 45, Biltmore, North Carolina.

Witham, E. C., Supt. of Schools, Putnam, Conn.

Woody, C., Director, Bureau of Mental Tests and Measurements, University of Michigan, Ann Arber, Michigan.

Woolley, Helen T., 264 Medbury Ave., Detroit, Michigan.

Worcester, D. A., Director, Bureau of Educational Measurements, Kansas State Normal, Emporia, Kansas.

Zirbes, Laura, Specialist in Reading, Teachers College, Columbia University, New York City.

JOURNAL *of* EDUCATIONAL RESEARCH

Volume VI SEPTEMBER, 1922 Number 2

ANOTHER EDUCATIONAL CAMPAIGN

J. W. Richardson
Director, Bureau of Research, Hibbing, Minnesota

In a former account of the use of campaign principles in elementary-school work, called "The Campaign Method in Elementary Education,"[1] we limited our discussion to progress in arithmetic under certain conditions. The conditions chosen were those which we recognize by the collective term of "a campaign." As stated in the above mentioned report, "A campaign is characterized: first, by a recognized need or desire on the part of those who organize it; second, by definite time limits, particularly for the close; third, by a definitely organized effort to stimulate and sustain interest. A fourth characteristic rests upon the fact that the success or failure of a campaign is determined by individuals other than those who organize it, whether it be by voters in the political compaign, converts in the religious revival, or subscribers to the fund. Furthermore, the compaign has a peculiar value in awakening the individual to the duties and privileges of co-operation whether civic, social, moral, or intellectual."

In making that report, we stated the belief that the plan would be just as feasible in the formal or mechanical phases of any elementary school work as in the four fundamental processes in whole numbers. We are now ready to state that the plan is certainly just as feasible in the subject of spelling as in arithmetic.

---

[1] Reported in the *Journal of Educational Research*, 2:481, June, 1920.

THE SPELLING CAMPAIGN

The Pupils—Twenty classes in fourth, fifth, and sixth grades, in two buildings, were each assigned to one of two groups. One of these groups was to be campaigned, and was designated as the experimental group; the other group, which was to take the initial and final tests, but was not to be campaigned, was called the control group. Each group consisted of ten classes, and both groups were estimated by principals to be of approximately equal ability.

The Material—Ten lists of fifty words each, with illustrative phrases, were used for tests. Each list was made up of the first fifty words in columns H to Q inclusive of the Ayres' Scale.

The Plan—Teachers were furnished a copy of the day's list on alternate mornings. They gave the test by reading the illustrative prase containing each word, and then repeating the word to be written in list form by the pupils.

The papers from the ten classes in the experimental group were corrected the day they were written and individual lists of misspelled words were returned to the teacher to be given to the pupils for study in the spelling period the next day. The pupils of this group were also told that the test would be repeated in exactly a month. No lists were returned nor announcements made to the classes in the control group.

TABLE I. PERCENT OF CORRECT SPELLINGS GRADE IV

AYRES' LIST	STANDARD	EXPERIMENTAL GROUP		CONTROL GROUP	
		First	Second	First	Second
H	98	96	98	95	96
I	96	96	99	91	93
J	94	96	98	92	94
K	92	94	97	84	86
L	88	90	93	81	84
M	84	83	94	73	78
N	79	81	93	67	71
O	73	73	91	58	66
P	66	63	88	46	44
Q	58	58	79	44	47
TOTAL.......		84	93	73	76

TABLE II. PERCENT OF CORRECT SPELLINGS GRADE V

Ayres' List	Standard	Experimental Group		Control Group	
		First	Second	First	Second
H	100	99	99	98	98
I	99	97	98	95	97
J	98	98	99	96	97
K	96	97	98	95	95
L	94	94	96	92	93
M	92	91	95	86	88
N	88	88	92	85	87
O	84	87	94	82	83
P	79	78	92	70	69
Q	73	75	89	65	68
Total.......		90	95	86	87

TABLE III. PERCENT OF CORRECT SPELLINGS GRADE VI

Ayres' List	Standard	Experimental Group		Control Group	
		First	Second	First	Second
H	100	99	99	99	99
I	100	98	99	97	98
J	100	98	99	98	99
K	99	98	99	97	97
L	98	98	98	97	98
M	96	95	98	93	94
N	94	93	96	91	92
O	92	91	97	89	89
P	88	85	94	83	86
Q	84	79	86	76	80
Total.......		93	97	92	93

It took a school month to give the ten tests on alternate days. The next month the tests were repeated in the same order (increasing difficulty, H to Q) and with same time intervals except that the first five were given on succeeding days instead of on alternating days, thus using three weeks for the second tests.

Each spelling period on no-test days was devoted to regular work by the control-group classes, and to study of individual misspellings by experimental-group classes. Mimeographed copies of word and phrase lists were also furnished pupils of the latter group to use if they wanted to review the entire lists.

Campaign Results—Table I shows the percent of correct spellings in each list of words made by the experimental and control groups in each trial. Tables II and III give the same information for grades v and vi, respectively.

Table IV is a summary of results. Part (a) gives the data for the experimental groups and part (b) the data for the control groups. A comparison of the two parts of the table shows the control groups to have been somewhat poorer spellers on the initial tests and that their gain was quite negligible in comparison with that of the experimental groups.

SUMMARY AND CONCLUSIONS

In the experimental groups the increase in percent of correct spellings of 33,400 words by the fourth-grade pupils was 9 (84 to 93); of 48,050 words by fifth-grade pupils was 5 (90 to 95); and of 30,050 words by sixth-grade pupils was 4 (93 to 97). These percent increases of 9, 5, and 4 were made in one-fifth of a school year in regular spelling periods yet the mean percent increase in the Ayres scale on the same lists for an entire year's work between fourth and fifth grades is 7.5 percent and between fifth and sixth grades is 4.8 percent.

TABLE IV. SUMMARY OF RESULTS

(*a*) *Experimental Groups*

GRADE	NUMBER OF PAPERS EACH TRIAL	NUMBER OF WORDS SPELLED EACH TRIAL	NUMBER OF WORDS MISSPELLED TRIAL		GAIN	
			First	Second	Number of Words	Per-cent
IV	668	33400	5527	2312	3215	58
V	961	48050	4552	2327	2225	49
VI	601	30050	12063	1045	1018	50
TOTAL	2230	111500	12142	5684	6458	53
PERCENT MISSPELLED..............			11	5	29[a]	

(*b*) *Control Groups*

GRADE	NUMBER OF PAPERS EACH TRIAL	NUMBER OF WORDS SPELLED EACH TRIAL	NUMBER OF WORDS MISSPELLED TRIAL		GAIN	
			First	Second	Number of Words	Per cent
IV	652	32600	8824	7913	911	10
V	772	38600	5346	4908	438	8
VI	821	41050	3288	2821	467	14
TOTAL	2245	112250	17458	15642	1816	10
PERCENT MISSPELLED...............			16	14	8[a]	

[a] Means number of words gained per pupil.

If pupils of the intermediate grades will eliminate 53 percent (Table IV) of their misspellings from the central half of the Ayres' list in one-fifth of a year's time simply because they know exactly where they stand, and that they will have a chance to prove the value of their efforts, are we not overlooking a great chance in the subjects, or divisions of a subject, which lend themselves to definite tests of development?

RECENT DEVELOPMENTS IN SILENT-READING TESTS

C. R. Stone
Principal, Gardenville School, St. Louis

Recent developments in the field of silent-reading tests have given us an equipment far superior to what we had two years ago. In particular, advancement has been made along three lines: (1) simplification in the testing procedure; (2) measurement of ability to read specific types of material; and (3) measurement of specific types of comprehension.

Part I. An Adequate Equipment of Silent-Reading Tests

Three types of silent-reading tests.—The most servicable reading tests and scales now available for group testing may be classified into three types as follows:

1. Short time paragraph tests with units of equal difficulty.
    a. Monroe Silent Reading Revised
        Test I for grades III, IV, and V (three forms)
        Test II for grades VI, VII, and VIII (three forms)
    b. The Burgess Picture Supplement Scales 1, 2, 3, and 4 for grades III to VIII inclusive.
2. Long time paragraph tests with units of increasing difficulty.
    a. The Thorndike-McCall Reading Scale (ten forms) for grades III to XII.
    b. The Haggerty Reading Examinations
        Sigma 1 for the lower grades (one form)
        Sigma 3, Part III, for grades VI to XII (two forms)
3. Tests for measuring rate and comprehension in rapid narrative-reading.
    a. Fordyce's Scale for Measuring Achievement in Reading
        Test No. 1 for grades III, IV, and V
        Test No. 2 for grades VI, VII, and VIII
    b. The Stone Series of Narrative-Reading Tests
        Test for grades III and IV
        Test for grades V and VI
        Test for the Junior High School

Types 1 and 2 are alike in that the pupil may re-read each paragraph unit until he thinks he can indicate a correct comprehension response. In this respect, all these tests measure a study type of comprehension ability. Those of Type 3 measure the rate of continuous reading and the ability of the pupil to grasp the main points in a single reading. The kind of reading material used and the character of the comprehension measured are radically different from that used in the tests of Types 1 and 2.

The Monroe Tests and the Burgess Scales.—While the Monroe Silent Reading Tests (Revised) and the Burgess Scales are alike in certain fundamental characteristics, as indicated in the preceding outline, they are different in certain respects. The paragraph units in the Burgess Scales are all of the same type of material, while in any one of the Monroe tests there are narrative, descriptive, informational, and poetical units. The Burgess plan is in accord with the present tendency in attempting to measure ability to read a specific type of material with a particular scale.

However, the Burgess scoring plan limits materially the usefulness of the scales. The Monroe scoring plan shows rate in terms of words read per minute and comprehension in terms of the number of correct responses. This makes it possible to graph the results for a school and easily to determine the relative efficiency as to rate and comprehension. In the Burgess test, the pupil's score depends upon the number of correct responses, and the score in no way reveals the pupil's relative rapidity and accuracy. Consequently, it is impossible to graph the results for a school and thus to show the relative efficiency as to speed and accuracy.

The Burgess scores range from 0 to 100 according to the number of paragraphs read successfully. The fourth grade must read successfully seven paragraphs to obtain the standard median score of 50, while the fifth grade is required to read eight paragraphs successfully to secure the standard median score of 50. The requirement is similarly increased for each succeeding grade. This plan has several disadvantages in comparison with the simpler plan used by Monroe—*i. e.,* that of allowing a comprehension score of one for each unit read successfully. The

Burgess plan complicates the graphing of the results for a school or system. It makes more difficult the locating of a pupil's grade level of reading ability. It also makes difficult the establishment of age norms.

After considerable experience with the Monroe and the Burgess tests, the writer believes that the usefulness of both to the teacher and to the supervisor would be materially increased if the scoring plan would show the number of units read as the speed measure and the percent of these units which are read successfully as the comprehension measure. Under the Monroe plan of scoring, the pupil or the school that has a high rate will almost always have a high comprehension; whereas if the comprehension accuracy were shown as a percent, we should find many cases of high rate and low accuracy. What the supervisor needs to know about the school and the class and what the teacher needs to know about the class and the individual pupil is the relative efficiency in rate and accuracy and the scores should reveal this without the necessity of referring to the test papers or to tedious comparisons with standards.

The Thorndike-McCall Scale and the Haggerty Reading Examinations.—About the same time that Burgess had concluded that difficulty or power tests are not applicable in measuring reading ability McCall had decided that the difficulty or power test is the best measure of reading ability. The units in the Thorndike-McCall Scale are progressive in difficulty. The first paragraph is relatively easy, the next somewhat more difficult, and so on. The paragraph sections of the Haggerty tests follow the same plan. The two differ, however, in the types of response required. Haggerty follows the plan of requiring the pupil to select from a number of listed responses. This plan lends itself to economy of time and to a high degree of accuracy in checking and scoring. The Thorndike-McCall plan is to have the pupil write out the answers to questions. This plan has certain disadvantages. The checking of the answers is time-consuming and there is danger of inaccuracy on the part of the grader. Furthermore, this plan limits the test very largely to a type of question that tests the pupil's ability to get certain definite detail, and consequently reduces to a minimum the use of questions that test the pupil's grasp of the major relational ideas or the essential total meaning of the paragraph.

The high cost of testing.—The Monroe and the Burgess tests are relatively inexpensive. The Thorndike-McCall tests are somewhat more expensive, and the Haggerty tests are very expensive. Testing material has been relatively more expensive than teaching material because testing material can be used but once while teaching material in the form of books and pamphlets can be used over and over. In the case of the Stone Series of Narrative-Reading Tests, while the initial cost is relatively high, the set of booklets and rate cards may be used over and over with incoming classes. Additional cost for answer sheets is very small.

Narrative-reading tests.—Besides measuring the ability of a pupil to do the study type of paragraph reading, the supervisor and the classroom teacher need to be able to measure his ability to grasp the main points in a narrative at a single reading and to be able to measure accurately the rate at which the whole narrative is read. The Stone Series of Narrative-Reading Tests[1] is designed for these purposes. At present there is available a test for the third and fourth grades, one for the fifth and sixth grades, and one for the junior high-school grades. Additional scales are in preparation. It is planned eventually to have one for each grade with two forms so that it will be possible to measure attainment and progress in any grade with a story of a difficulty level of that grade. This means that the teacher of each grade will have a testing equipment on narrative reading that will enable her to diagnose her incoming class and to measure progress after a certain period of instruction. This regular use of testing equipment the same as we regularly use the reading book, the writer conceives to be essential to the most effective type of teaching.

Part II will give a detailed account of the construction and the use of the Narrative Reading Test for the Junior High School.

---

[1] Published by The Public School Publishing Co., Bloomington, Ill.

PART II. THE STONE SERIES OF NARRATIVE-READING TESTS

Difficulties involved in constructing a measure of narrative-reading ability.—Recent reading tests have been paragraph tests. It is much easier to measure ability to read paragraphs than ability to read a complete narrative. It is very easy to adopt a simple procedure in a paragraph test, while it is almost impossible to do so in a narrative test. In a speed test in paragraph reading, in such tests as the Monroe or the Burgess, the time taken to do the testing is very short, while it is almost impossible to give a narrative test in less time than it takes to give a power or difficulty test in paragraph reading.

The problem of obtaining an accurate measure of rate in some such terms as the number of words per minute has been a serious one in devising narrative tests. Gray invented a very accurate means of finding the pupil's rate of reading, but his plan involves individual testing; and this is too time-consuming for regular classroom purposes. On the other hand, classroom teachers and supervisors have become very skeptical of the commonly used plan of having the pupils read a minute or so and at a signal mark the last word read.

Another problem is so to relate the rate and comprehension tests that we may have a rate measure and a comprehension measure based on the same reading act. Only when this is done can we satisfactorily characterize the reader as slow and accurate, rapid and accurate, slow and inaccurate, or rapid and inaccurate. No group narrative test printed previous to 1922 solves this problem.

Construction of the Junior High School Test.—In the construction of the Junior High School Test a number of preliminary experiments were tried, an account of which it is impossible to give in this brief article. The following principles were kept in mind from the beginning:

a. The rate test should include for all pupils the reading of the whole narrative.
b. The test for the upper grades should be based upon a narrative of upper-grade level.

c. The comprehension test should measure the pupil's grasp of the leading points of the narrative.

d. Both the rate test and the comprehension test should be as simple in the testing procedure as possible.

e. The narrative should be of such length that no pupil would complete it in less than two minutes.

The first story tried was too easy; but a passage from Cooper's "The Spy" proved to be of eighth-grade level, and it was adopted. This narrative passage is about 1400 words in length. During the spring of 1921, eighth-grade pupils to the number of 563 were tested. The median rate was 234 words read per minute. On the comprehension test of ten questions, the median was 64, ten points being allowed for each question correctly answered. This extensive trial of the test followed preliminary trials and revisions. By this time the questions had been found to be approximately equal in difficulty. However, in constructing diagnostic charts for interpreting results it was found that ten questions did not give a sufficiently refined scale.

Very early in the work of constructing the test it was decided to have preparatory rate and comprehension tests precede the real tests. The preparatory tests are carried out in all details like the real tests. The unit of reading is, of course, much shorter.

The rate test is conducted as follows. The pupils are told that they are to read the narrative through from beginning to end, reading as rapidly as they can and at the same time get the story. Each pupil understands that he is not to look up until he has finished reading the story. Then he looks up and gets his rate from a card then being exhibited and copies it on his paper. A new rate card is exhibited each five seconds. The number on the card represents the number of words read per minute. A pupil, or some other assistant who has been trained for the purpose, exhibits the rate cards under the supervision of the examiner.

After the pupil has recorded his rate score, he proceeds immediately to answer the questions without being permitted to re-read. Each question has five answers listed underneath it, only one of which is correct. The pupil answers the question by selecting one of the answers and indicating it on his answer paper. There are twenty questions in the tests as published. He is allowed as much time as he needs in answering the questions; he is to take out a book and read until all have finished.

It takes about an hour to give the preparatory and the real tests. This, however, includes the checking and rechecking of the answers by the pupils—a procedure which makes it unnecessary for the teacher to spend time outside of school hours in scoring the papers.

Results of testing seventh- and eighth-grade pupils: mid-year, mid-grade.—Figure 1 shows the distribution of the rate scores of 320 eighth-grade pupils on a scale ranging from 500 to 100 words per minute. The graph shows that there were 15 pupils who read at a rate above 500. Only 1 pupil read at a rate as low as 100 words per minute. The mode is 220 and the median 245. We see that the curve is skewed somewhat to the lower end of the scale. Otherwise the distribution corresponds fairly closely to the normal curve of distribution.

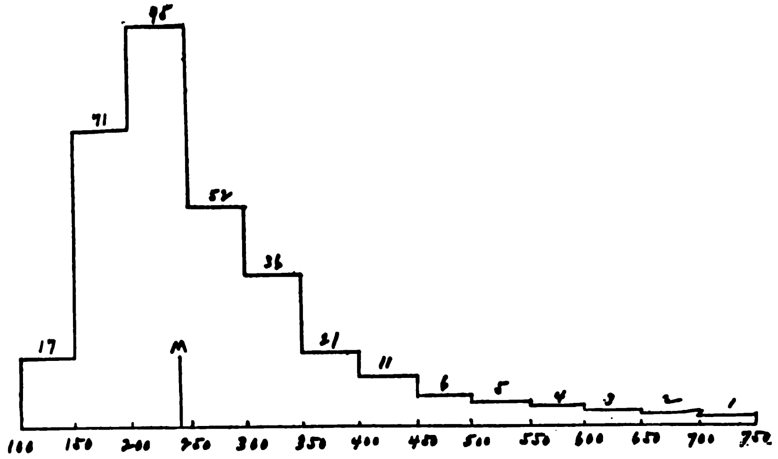


FIGURE 1. THE DISTRIBUTION OF RATE SCORES OF 320 EIGHTH-GRADE PUPILS

The usual wide range of attainment is clearly in evidence. The problem arises as to whether there is really such wide variability in the rate attainments of eighth-grade pupils or whether the marked variability is due to inherent defects in the measuring instrument. For instance, is the story of uniform interest to boys and girls alike, or do some of the pupils read rapidly because they are gripped by the story while others read slowly because the story does not appeal to them? It was the general opinion of those connected with the testing that the story is generally interesting to the pupils and about equally so to both boys and girls. However, a comparatively small number of pupils stated after the test that they did not care for the story.

It may be remarked in passing that in selecting the stories for the fifth- and sixth-grade test, preliminary experiments were carried out to aid in securing stories having a uniformly strong appeal to the pupils. Six stories were selected with a view of getting those that would be interesting to all pupils. All the pupils in a given room were asked to read six stories at one sitting and to put down afterwards the order of preference. The two stories having the highest degree of uniformity of appeal were selected for the test. The rate score and the comprehension score, in each case, is the average of the scores made on the two stories. This plan tends to decrease somewhat the wide variability in the pupils' scores.

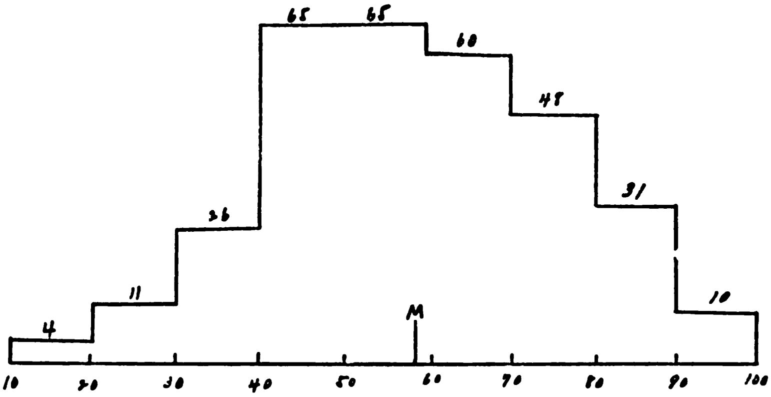


FIGURE 2. THE DISTRIBUTION OF THE COMPREHENSION SCORES OF 320 EIGHTH-GRADE PUPILS

Figure 2 shows the distribution of comprehension scores for the Junior High School Test on a scale ranging from 0 to 100, counting 5 for each of the 20 questions. The distribution curve follows rather closely the normal curve of distribution. It will be noted that no pupil made a perfect score, or in other words, that no pupil answered all 20 questions correctly. Likewise, no pupil made 0. Consequently all pupils were measured. The median is 57 or 11.4 questions answered correctly. Although it is considered that this story is of about eighth-grade level, the results for 274 seventh-grade pupils show that it is usable in the seventh grade also. The comparative results of the seventh and eighth grades indicate that the test will also be an accurate measuring instrument for the ninth grade.

In deriving the medians for Table I, data from only those schools which gave the test in both the mid-seventh grade and the mid-eighth grade were included. This was to insure more accurate rate and comprehension intervals between the seventh and the eighth grades.

TABLE I. MID-YEAR, MID-GRADE MEDIANS FOR ST. LOUIS

GRADE	RATE	COMPREHENSION	NO. PUPILS
VII	218	10.2	274
VIII......	245	11.4	320

The use of the diagnostic charts.—To the classroom teacher, the most important fact that a test of this sort should reveal is the type of reading each pupil does. In almost any room or class of pupils, we find the slow accurate reader, the rapid accurate reader, the rapid inaccurate reader, the slow inaccurate reader, and all of the intermediate combinations of the two elements of rate and accuracy. The diagnostic chart is designed to show the teacher graphically just how her class distributes among nine rate and comprehension groups. The chart is constructed in such a way that on the basis of the 320 eighth-grade pupils tested approximately 50 percent will be in the middle comprehension division, about 25 percent in the inferior division, and about 25 percent in division labeled good. The same plan is followed with reference to the rate divisions.

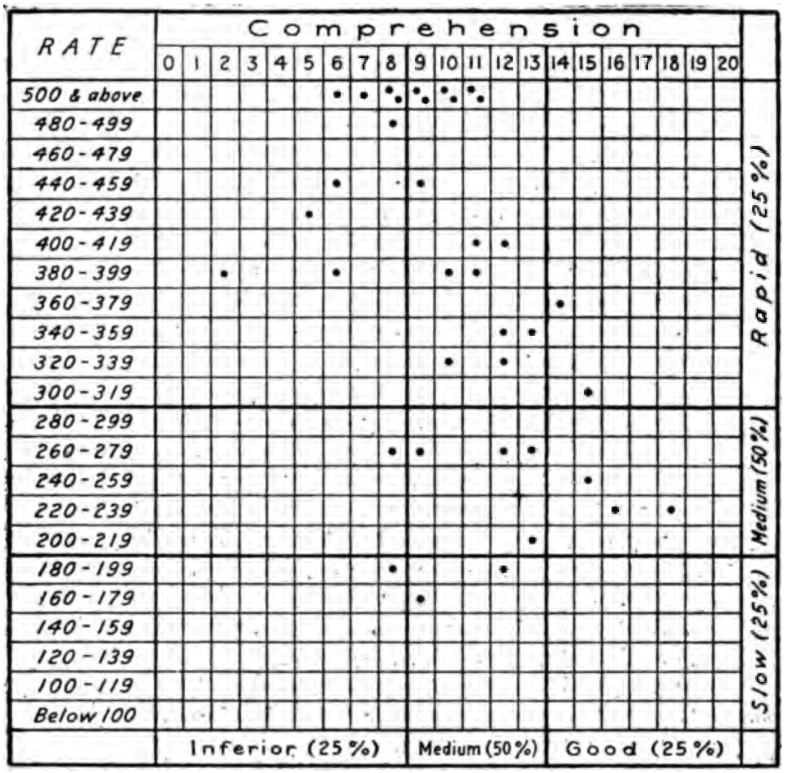


FIGURE 3. THE DIAGNOSTIC CHART FOR A MID-EIGHTH GRADE CLASS IN SCHOOL A. 37 PUPILS

Charts for three eighth-grade classes are reproduced representing three distinct types of classes with reference to reading efficiency. Figure 3 is the chart for a class in School A that is far above the average in rate and somewhat below the average in comprehension. The distribution indicates a small negative correlation between rate and comprehension. The fact is that this class in its training had been speeded up at the expense of comprehension. A number of these pupils should be told to read somewhat more slowly. This point should be particularly emphasized with the nine pupils falling in the rapid-inferior division. It will be noted that in this class the two pupils ranking highest in comprehension are medium-rate readers. The medians in rate and comprehension for this class were 390 and 10.6 respectively. These may be compared with the tentative city medians of 245 and 11.4.

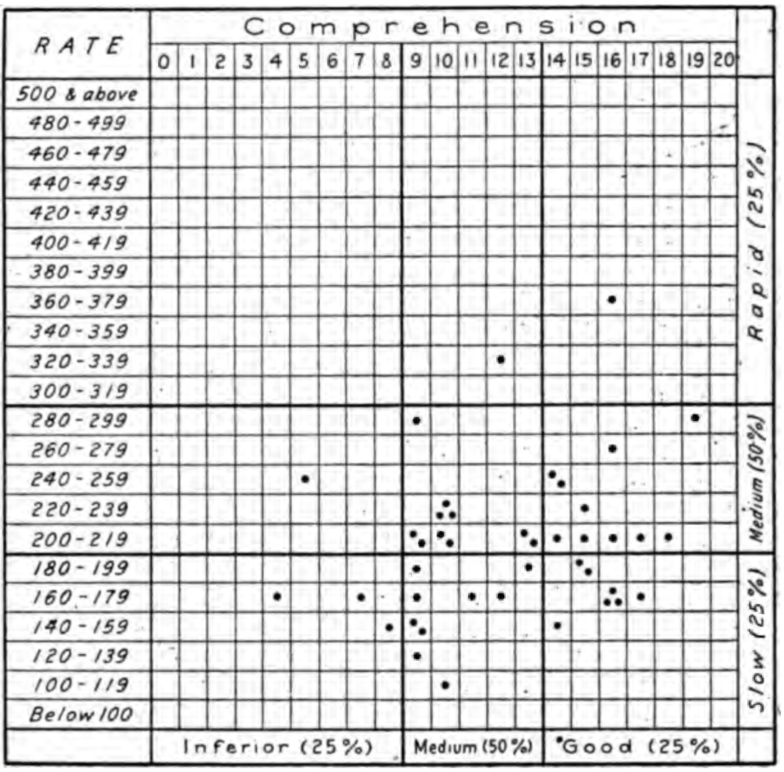


FIGURE 4. THE DIAGNOSTIC CHART FOR A MID-EIGHTH GRADE CLASS IN SCHOOL B. 42 PUPILS

The next chart, Figure 4, reveals a situation in School B that is radically different than that in School A in the same grade. This class is considerably above the standard in comprehension (median 13.0) and considerably below in rate (median 202). In the preceding chart the dots tended to distribute toward the upper left-hand corner, while in this chart they tend to fall toward the lower right-hand corner. There are seven pupils who are slow but good in comprehension. It is possible that this school has stressed thoroughness at the expense of speed. It appears that the training now needed by this class is of an opposite type from that required of School A.

It is not to be expected, of course, that each pupil or each class will perform according to some standard balance of speed and comprehension. The most efficient speed for one pupil may not be the best speed for another. The type of mind of the pupil will be an important factor. School B is in a community of relatively low social status. The intelligence of the pupils is below standard. The writer's experience leads him to believe that the pupil of low intelligence, in order to comprehend equally as well as the bright pupil, must read more slowly than the bright pupil. It is not meant by this that the slower one reads the better he comprehends. It is possible that those who are slow and inferior would comprehend better by reading faster. The problem is for the pupil to find the rate at which he comprehends best and then gradually to increase his rate as his ability to comprehend develops.

The class in School C, Figure 5, shows a greater variety in types of readers than either of the other two. The medians (234 and 10.0) are somewhat below the tentative norms. There are more pupils falling in the slow-inferior group than in either of the other two classes. Evidently these slow-inferior pupils are working at a great disadvantage if placed in the same class and given the same material as the pupils who are rapid or medium in rate and good in comprehension. The pupils of this class should be grouped according to their needs and should be given reading material of different levels as well as different types of training and experience.[2] Most of the pupils of this class should have training for developing speed.

Relation of rate and comprehension.—The teacher's manual for a set of readers planned for considerable amounts of silent reading contains the following statement: "The fastest readers usually make the best readers. Therefore, if you develop speed, you improve the quality of the child's reading." This reasoning not only does not accord with the results of research work in reading but it also fails to accord with the common experience

---

[2] For specific details see Stone, Clarence R., *Silent and oral reading.* Boston: Houghton Mifflin Company, 1922. 295 pp.

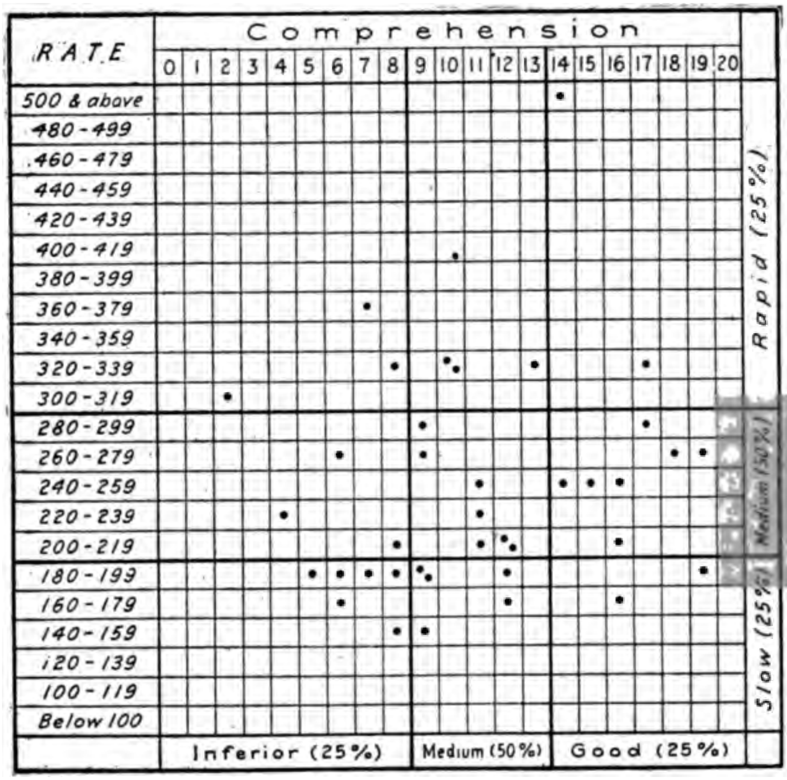


FIGURE 5. THE DIAGNOSTIC CHART FOR A MID-EIGHTH GRADE CLASS IN SCHOOL C. 38 PUPILS

of teachers. The data resulting from the application of the Junior High School Test show about what the St. Louis and the Cleveland Surveys showed, namely that the rapid-good pupils and the slow-poor pupils are about twice as numerous as the rapid-poor and the slow-good. However, this does not mean that increased speed brings improved quality. There is reason to believe that School A has developed speed at the expense of quality, while School B has not given sufficient attention to speed. No doubt very slow readers may comprehend better if they read somewhat faster. Probably all such pupils should be encouraged to speed up and should be given types of drill designed to produce a wider span of recognition per eye-pause.

On the other hand, many pupils have reached the point in the speed scale that for them gives the maximum efficiency in comprehension. For them to be speeded up without a corresponding development in ability to comprehend is likely to lower the quality of their reading. There is also the rapid-careless reader who should be given training that develops accuracy. A decrease in his speed will probably be advisable.

Pupils should be grouped and given reading instruction according to their needs. The diagnostic chart is designed to assist in determining their needs on the basis of each pupil's relative attainment in rate and comprehension.

Summary.—Each of the narrative reading tests has the following distinguishing characteristics:

1. There is a preparatory test, shorter than the real test, but carried out in every detail just like the real test.

2. There is a unique plan for measuring the rate involving the use of a set of rate cards. To get his rate score all the pupil is required to do is to look up when he comes to the end of the story and copy the number then being exhibited.

3. The comprehension test is upon the reading which has just been measured as to rate. It consists of twenty questions testing the pupil's grasp of the leading points of the story. The pupil selects one of the five listed answers for each question. The plan is simple and economical in time; it is an accurate measure of the comprehension of a narrative.

4. The tests are concerned with a specific type of reading material and they measure a specific type of comprehension.

5. Since the pupil enters his data on an inexpensive sheet, the test material may be used over and over again, thereby reducing to a minimum the per pupil cost of testing.

6. By using the diagnostic chart, the teacher can easily make practical use of the results in analyzing individual cases and in grouping her pupils for more effective instruction in accordance with individual needs.

7. The time required for testing a group of pupils including the checking of the answers and the recording of the scores is from 40 to 60 minutes. This means that the teacher is not required to do a lot of checking and scoring outside of school time.

SOME ESSENTIALS IN SCHOOL SUPERVISION[1]

E. J. ASHBAUGH
Ohio State University

I suppose I may assume general agreement to the proposition that supervision is for the purpose of improving classroom instruction. If the supervisory force does not function through the teacher in increased efficiency in class work, its existence cannot be justified nor maintained. This increased efficiency must manifest itself through either a greater quantity of work of a given quality, an equal quantity of better quality or an equal quantity and quality at a lesser time expenditure. The supervisory group has tended to be recruited either through special training for this type of work or through selection of teachers who had had marked success in some particular line and who, it was therefore thought, might have ability to direct other teachers to acquire similar success. But whatever may have been the training or the method of selection, there are certain essentials which must be borne in mind if the work is to be satisfactory. It is the purpose of this paper to discuss a few of these.

The selection of the appropriate subjects by means of which to attain desirable educational objectives is a problem in itself. Tradition, current practice, and legal or institutional requirements are often more potent than the supervisory force which must deal with these subjects in the schools. However, an educational philosophy which interprets the content of the subjects in terms of present day needs and probable future needs may do much in determining the educational ends to be secured by means of these subjects.

DEFINITE GOALS

Assuming that these general subject fields are already determined, it seems to me that the first and most important essential in supervision is a rather accurate knowledge of definite goals of attainment. The best supervisors are not content today with intangible goals, ethereal objectives, indefinite generalities. Teaching reading in order that children may become cultured individuals is not a satisfactory statement for the supervisor, however true it may be as an ultimate objective. The supervisor must not only have an ultimate end in view but must have in mind a large number of intermediate stations definitely charted as guides in reaching this ultimate. I like to think of these intermediate goals as towns or cities along a well-marked automobile road. Each is definitely marked so we know when we have reached it. We know how far we have come since leaving the last place; we know what the next one is on the route and how far we must travel before arriving there.

---

[1] An address presented at the Educational Research Department of the Ohio State Teachers Association, Cedar Point, Ohio, June 28, 1922.

Confining ourselves for a bit to the field already mentioned, to reading, let us seek farther for acceptable objectives. One says that we desire our children to learn to read in order to get the thought from the printed page. This is better than the former statement of aim, but it is still unacceptable from the standpoint of supervision. It is not a specific goal. Continuing the analogy to an automobile trip, the first statement (*i.e.*, of culture as an objective) would be like saying that the place to go is where the climate is pleasant, the people hospitable, and the roads good the year round. Such a place may indeed be desirable but the tourist who is directed in his journey only by such a statement will certainly be handicapped in proceeding on his journey. The second statement, namely as to getting thought from the printed page, might be likened in the same analogy to saying that the place to go is to the city in Missouri where there is satisfactory hotel and garage service at the least possible price. If that were the goal determined upon, it might serve to start the tourist off in the right general direction; but how would he know when he got there? The state of "thought getting" is nearly as large as the state of Missouri.

Further, in going to Missouri one might take any number of days largely at his pleasure; but in guiding the child to the desired goal in reading, one may not loiter, for the length of time he may remain in school to profit by our guidance is problematical. It is certain that unless he reaches this goal in not more than four or five years, he will be greatly handicapped throughout the remainder of his career. Hence it is even more important that the goal be defined definitely and that landmarks be set along the way; for progress may then be noted and greater assurance may be had that the ultimate objective will be reached before it is too late.

A goal more specific than "somewhere in Missouri" would be Marysville, Mo. Similarly, a goal in reading more specific than "ability to get thought from the printed page" is "ability to read ordinary or non-technical material at the rate of not less than 240 words per minute and at the same time to get the thought so as to be able to reproduce, by answering questions or otherwise, not less than 95 percent of the ideas expressed in the selection."

This is a definite goal which may be set for attainment by at least half of the children by the end of the eighth grade. Having established such a goal, we may look for the intermediate goals or guides which will enable us to know what progress we are making toward this more remote objective. We may, for example, set up grade goals as follows:

GRADE	SPEED	COMPREHENSION (percent)
I	50	50
II	95	60
III	125	70
IV	150	80
V	180	85
VI	200	90
VII	220	93
VIII	240	95

These goals, it is to be understood, are hypothetical and may be at considerable variance with those which should be established, although they are not greatly different from the norms of attainment which have been ascertained by workers in this field. Only tentatively, however, should such norms be considered goals, only, indeed, until the necessary analytical and experimental work has been done to verify them or to establish new ones.

A vast amount of such analytical and experimental work is needed. In order to set a satisfactory goal for the end of the school course, such questions as the following need to be answered: At what rate and with what degree of comprehension will the average adult twenty-five or thirty years hence need to be able to read in order to live successfully his full measure of life as a worthy member of society at that time? We must set our goal in terms of the future. The school must prepare the child for adult life because it will be with those, who like himself are now enjoying greater educational advantages than their parents, that he must live and work. Doubtless this future need must be determined by estimate from the ascertained needs of the present. Even these, however, are thus far undetermined.

Other questions that will need to be answered are: What degree of overlearning must be accomplished in order that the desired residue may continue to exist after drill days are over? What adjustment of goals should be made because of variation in type of material? What other goals are legitimate objectives in reading besides speed and comprehension? What are definite goals in appreciation of literature, such as appropriateness of style, vividness of description, aptness of illustration, etc.

Or, to illustrate what I mean from another field, the supervision of the teaching of arithmetic cannot be done well with only general objectives. Here again the goals must be definite and a large number of intermediate points should be definitely charted. Teaching children to add well is vague until *how well* is indicated in terms of rate and accuracy in respect to a definite type of material. Instructing a teacher to develop ability to use common and decimal fractions without defining the ability more specifically leaves large leeway for the teacher either to be satisfied with too little ability or to insist upon the acquisition of ability far beyond any need which is likely to arise in the life of either child or adult.

What has been shown for reading and arithmetic is quite as true in the other subject-matter fields. Lack of definite goals, clearly indicated by the supervisor, is readily recognized in the variation among the teachers of a system as to what constitutes satisfactory work on the part of the pupils.

ATTAINING THE OBJECTIVES

The second essential follows logically from the first, namely, a rather definite knowledge of how to attain these objectives. Supervisors are, in theory at least, sources of help and inspiration for the teachers in the accomplishment of the tasks to be done. There has been much written and more said about the danger of over supervision, of the injustice in crushing out the teacher's initiative, of supervision making the teacher only a cog in the machine, etc. True supervision does none of these things, though it may prevent the teacher's initiative from straying far from the desired goals of her educational work.

True supervision certainly does not mean that the supervisor tells the teacher or demonstrates to her every new step along the way. In many of our schools the so-called penmanship supervisor does all the penmanship teaching that is done; the art supervisor does all the art teaching; the music supervisor all the music teaching. These are not supervisors; they are teachers assigned to a special subject in all rooms instead of to a group of subjects in one room. The true supervisor, operating directly upon the teacher and through the medium of the course of study, sets the definite goals in terms of quality and quantity of achievement. Page terminals in the textbook may indicate something of quantity, but they never define quality of work. Let us rule them out and substitute real worthwhile goals which define both quantity and quality.

Having defined the goal in the subject for the grade and perhaps indicated some of the landmarks along the way, the true supervisor will give the teacher an opportunity to exercise her initiative in attaining this goal. Detroit has been trying out a scheme whereby the supervisor is prohibited from visiting the classroom or making any suggestion to the teacher except upon request. This not only gives the teacher an opportunity to exercise her initiative but it also places responsibility squarely upon her either to attain the goal alone or to request help. If she succeeds alone, the glory is hers and the supervisor has been left free to work where needed. If she realizes she needs help to attain this goal and asks for it, she is in the best condition to profit by the help the supervisor can give and hence the time of the supervisor is economized. If she does not call for help and fails to attain the objective set, the responsibility rests squarely upon her and indicates that either she can not judge accurately the progress of her pupils or over-estimates her own ability or refuses to acknowledge her need. In either of these cases she certainly needs supervision.

Regardless of whether the responsibility for recognition of needs be placed upon teacher or supervisor, whenever the need exists, the supervisor should be able to furnish it. Just how this aid should be given may vary in terms of the personality of supervisor, teacher, type of need, etc. The supervisor may demonstrate, taking the class before the teacher and illustrating the working procedure by which the desired results may be obtained economically and efficiently. She may, in personal private conference with the teacher, tell how the teaching process should be carried on or give suggestions which will supplement the ideas which the teacher has already been utilizing. She may, either in conference or by letter, direct the teacher's attention to available literature in which the teacher will find specific help in her difficulty. The true supervisor may use any or all of these means to secure the accomplishment of the goal but in all cases she must know fairly definitely how to attain the goal or where the information may be found.

CHECKING ATTAINMENT

The third essential is closely associated with the first two. Efficient supervision is dependent upon the supervisor's knowledge and use of valid means for checking the extent to which the goals set forth have been attained. To set forth ends to be secured without ascertaining whether or not they have been attained is folly. It is worse than folly, since it will quickly break down the effort on the part of the teachers to attain these objectives. The teacher who questions the validity of the objective or who has ends of her own which she considers more important will quickly cease to strive for the former when she learns that no check is made of her attainment of the objective in question.

Probably the best reason for the use of adequate means of checking attainment is exemplified in the old adage, "Nothing succeeds like success." Children and teacher alike wish to know of the success of their efforts. If the objective for eighth-grade reading has been set at 240 words per minute with 95 percent comprehension, every eighth-grade child and teacher will be interested in knowing whether that goal has been reached or exceeded. Every child who has passed it will feel the glow of successful effort. This will be shared by his teacher. The placing of goals upon an objective basis removes the personal prejudice charge on the part of those who fail. Supervisors may well expend a part of their time and ingenuity in devising objective tests for measurement of success in attaining their definite goals, and they must certainly utilize those tests which are already available.

But it is more important that progress be known before the end of the year than at that time. Diagnostic tests which not only show status attained but also reveal weaknesses are of supreme value in supervisory work. To the extent that difficulties are discovered early, and to the extent that no time is lost in futile effort and no wrong habits are formed, progress will approach maximum rate. The teacher should be keenly alive to these facts and ready to use the means available to check on these items. But it is the function of supervision as it is organized in most systems to see that these items are cared for whether the teacher is alive to their importance or not. Hence the supervisor must be aware of and use these instruments for checking progress.

Supervision without definite goals and objective means of checking progress is likely to be of little value. A state wide study of handwriting in Iowa a few years ago shed a little light upon this phase of the topic under discussion. Rural schools in that state have practically no supervision, since one county superintendent for 150 to 225 one-room schools scattered over an area of 500 to 900 square miles can not possibly supervise much, even if he had nothing else to do. And he certainly has many other duties. On the other hand, all cities of 10,000 or more had, and had had for a number of years, special supervisors of this subject. A single system of penmanship instruction was in use in most of these cities. But definite goals of attainment in terms of specific rates of speed and qualities of form were not in use and objective means of checking attainment were also lacking. When, therefore, the samples of writing were carefully rated by the Ayres scale, the results revealed that the median quality of handwriting in all grades in city schools averaged less than one-half step on the scale better than the quality of writing of the rural children and in the third and fifth grades the quality of writing of rural children was actually better than that of city children. A similar survey in the fundamentals of arithmetic showed an advantage for city children of only about one-half the average inter-grade advance. Such lack of significant difference between the achievement of supervised and non-supervised schools points clearly to ineffective supervision.

OPPORTUNITIES FOR RESEARCH

These three essentials call for much research work and supervisors in daily contact with both teachers and children are in a particularly strategic position to conduct at least certain phases of it. The definite goals by subject and grade can best be determined in the schoolroom. Supervisors may well set the goal high and then utilize all available knowledge in method and psychology of learning to set new records of accomplishment. The little that has been done has shown conclusively that the progress that is normally made in rate of reading in two years, can be attained in less than half that time with no loss of degree of comprehension. Richardson's experiments in spelling and arithmetic have shown that "campaign" methods enable one to speed up the process of attaining specific goals very materially. Analysis of social needs such as the work of Horn in spelling and the social sciences and of Wilson in arithmetic are fundamental to the determination of the proper content of our curriculum. Such work as that of Charters in language errors lays the foundation for specific application of corrective measures. Every supervisor could add to the knowledge now available if she would secure the cooperation of her teachers and their pupils in a small but definite bit of research work each year.

This naturally suggests another essential in supervision, namely the ability of the supervisor to secure the whole-hearted cooperation of her teachers. This is secured only through the demonstration of a sympathetic, helpful attitude. Whenever a teacher discovers that the supervisor is a sympathetic source of help and inspiration, her cooperation is at once guaranteed. The goal set up by the supervisor may be new and apparently unattainable, but she will give her utmost effort knowing that the supervisor will gladly lend a helping hand and will either show her how to attain the proposed objective, or cite her to the literature which will explain. The devices suggested may be strange but they will receive consideration because she will know that the supervisor is as vitally interested in her success as in the success of these boys and girls. She will also know that a meed of praise will be hers when the checking is done and when it is seen that aspiring to the goals and following both her own best initiative and the advice of her supervisor the desired results have been attained.

The program suggested here is applicable in any place where there is any supervision at all. If you are a superintendent in a small place where all supervision must be done by you, it is still possible for you to follow it. If you feel you can do more, select one subject; make yourself acquainted with the literature in the field so that you may set up definite goals of attainment which will be in line with the best that is known on the subject, be able to cite your teacher who needs help to books and articles which will aid her in attaining these goals; use the means available to check results; and secure the sympathetic cooperation of your teachers in doing a bit of real research on this problem in your own schools. It means some time and effort and doubtless you feel that your time and effort are about all taken now. But you owe it to your school, to your teachers, and to yourself to do this. It means service of the best type to both school and teachers, and it means professional development of yourself at the same time.

CAN TEACHERS SELECT BRIGHT AND DULL PUPILS?

G. F. Varner
College of Education, University of Minnesota

Scientific investigation of the use of rating scales has shown that ratings are not reliable measures of human traits. Rugg,[1] in discussing the rating of officers in the United Staes Army, states that: "It was very improbable that an officer was located within even his proper 'fifth' of the entire scale by an 'official' rating." While these findings discredit teachers' ratings of pupils' abilities it might be possible that teachers can select those pupils who represent the two extremes of a trait. This paper is a discussion of an attempt to measure the ability of teachers to select the extremes of general intelligence. Intelligence was chosen in preference to other traits because: (1) It is the only trait for which satisfactory objective measures are available with which to check the ratings. (2) Teachers' estimates are often used in the selection of bright and dull pupils to be put into special classes. (3) If we believe that the bright children should be given a chance to progress more rapidly than the average children, then it becomes important to know whether the teachers can, without the use of intelligence tests, select those who might profit by such an opportunity. (4) Unless teachers can select the bright pupils, they will not know how to advise pupils intelligently in the selection of a vocation. Unless a teacher is able to estimate which are the bright children, some children will be advised to enter professional occupations when their ability is not great enough to warrant their preparation for such a life work.

From the data presented here an attempt will be made to answer the following questions: (1) How reliable are teachers' selections of bright and dull pupils? (2) Is it more difficult to select the bright pupils than it is to select the dull pupils?

---

[1] Rugg, Harold Q., "Is the rating of human character practicable?" *Journal of Educational Psychology*, 12:436, November, 1921.

(3) Does it become more or less difficult to select bright and dull pupils as they become older?

In the schools of St. Paul, Minnesota, the teachers select the upper and the lower twenty percents of their pupils in general intelligence in order to supplement the intelligence test scores in determining which pupils will be put in special classes. The material used in this investigation was obtained from the St. Paul schools in 1920. It concerns six eighth-grade classes containing 206 pupils and three second-grade classes containing 80 pupils. The teachers were given a definition of intelligence and were directed to select the best and the poorest twenty percents in their grades. The eighth-grade pupils were tested with the Haggerty Intelligence Examination, Delta 2 and the second grades were tested both with the Haggerty Examination, Delta 1 and with the Pressey Primer Scale. One of the second grades was not tested with the Delta 1, so in this case the results were for the Pressey test only.

The scores of all the pupils were translated into intelligence quotients and the teachers' selections were compared with the intelligence quotients, as well as with the actual test scores. It was found that the teachers' selections compared more favorably with the intelligence quotients than with the actual scores. So for this report the intelligence quotients are taken as affording the mental classification to use and the results are reported for the intelligence quotients only. Since the teachers were directed to select the highest and the lowest pupils in their grades regardless of the general intelligence level of their pupils, the comparisons must be made for each grade separately. The comparisons were made by dividing all of the pupils of each grade into twenty percent classes according to their intelligence quotients. The bright and dull pupils as selected by the teachers could then be located within the twenty percent class to which they belonged according to the test classification. The results of this comparison are shown in Table I for the eighth-grade classes and in Table II for the second-grade classes. They are summarized by percents in Tables III and IV.

TABLE I. A COMPARISON OF THE UPPER AND LOWER TWENTY PERCENTS OF PUPILS AS SELECTED BY THE TEACHERS WITH THE INTELLIGENCE QUOTIENTS AS FOUND FROM THE HAGGERTY DELTA 2 INTELLIGENCE EXAMINATION. (EIGHTH GRADE.)

School	Selected by Teachers	Classification by Intelligence Quotients				
		First 20 percent highest	Second 20 percent	Third 20 percent	Fourth 20 percent	Fifth 20 percent lowest
I.	Upper 20 percent..	4	2		1	1
	Lower 20 percent.			1	1	5
II.	Upper 20 percent..	4	2	2		
	Lower 20 percent..				3	3
III.	Upper 20 percent..	1	1	2	3	
	Lower 20 percent..				2	5
IV.	Upper 20 percent..	3	2	3		
	Lower 20 percent..		1		1	3
V.	Upper 20 percent..	4	2	1		
	Lower 20 percent..				1	3
VI.	Upper 20 percent..	4	3	1	1	
	Lower 20 percent..			2	2	5
Totals	Upper 20 percent..	20	12	9	5	1
	Lower 20 percent.		1	3	10	24
Percents	Upper 20 percent..	42.55	25.53	19.15	10.63	2.13
	Lower 20 percent..		2.63	7.90	26.32	63.17

TABLE II. A COMPARISON OF THE UPPER AND LOWER TWENTY PERCENTS OF PUPILS AS SELECTED BY THE TEACHERS WITH THE INTELLIGENCE QUOTIENTS AS FOUND FROM THE HAGGERTY DELTA I AND PRESSEY INTELLIGENCE EXAMINATIONS. (SECOND GRADE.)

School	Selected by Teachers	Classification by Intelligence Quotients				
		First 20 percent highest	Second 20 percent	Third 20 percent	Fourth 20 percent	Fifth 20 percent lowest
I.	Upper 20 percent..	1	2	2	4	1
	Lower 20 percent..		2	2	1	4
II.	Upper 20 percent..	2	1	1	1	
	Lower 20 percent..			1		4
III.	Upper 20 percent..	1	1		1	
	Lower 20 percent..		1	1		1
Totals	Upper 20 percent..	4	4	3	6	1
	Lower 20 percent..		3	4	1	9
Percents	Upper 20 percent..	22.22	22.22	16.65	33.33	5.55
	Lower 20 percent..		17.24	23.53	5.83	52.94

Table I shows, for example, that of the eight pupils selected by the teacher of School I to be in the upper twenty percent class in intelligence, four were put in that class by the intelligence quotient classification, two were put in the second class, one was put in the fourth class and one was put in the fifth or lowest class. Of the seven pupils selected by the teacher to be in the lowest twenty percent class in intelligence, five were classified in the same class by the intelligence quotients, one in the class above and one in the middle class. Table II gives the same data for the second-grade classes.

TABLE III. A COMPARISON OF THE ABILITY OF TEACHERS TO SELECT BRIGHT AND DULL PUPILS. (EIGHTH GRADE.)

	Percent of Pupils in Same or Different Classes According to Intelligence Tests				
	Same class	One class removed	Two classes removed	Three classes removed	Four classes removed
Upper twenty percent according to teachers.........	42.6	25.5	19.2	10.6	2.1
Lower twenty percent according to teachers.........	63.2	26.3	7.9	2.6	0.0

TABLE IV. A COMPARISON OF THE ABILITY OF TEACHERS TO SELECT BRIGHT AND DULL PUPILS. (SECOND GRADE.)

	Percent of Pupils in Same or Different Classes According to Intelligence Tests				
	Same class	One class removed	Two classes removed	Three classes removed	Four classes removed
Upper twenty percent according to teachers.........	22.2	22.2	16.7	33.3	5.6
Lower twenty percent according to teachers.........	52.9	5.9	23.5	17.2	0.0

Tables III and IV are summaries of Tables I and II. They show the percent of pupils concerning whom the teachers' estimates agreed with the test scores. They also show the percent of pupils concerning whom the teachers disagreed with the test results and the amount of the disagreement. Table III shows that of those selected by the teachers to be in the upper 20 percent of all the eighth-grade classes, about 43 percent were likewise placed in the upper 20 percent of these classes according to their intelligence quotients. If we may assume that the test results are correct, 25.5 percent of the pupils were misrated by the teachers to the extent of one class, 19.2 percent to the extent of two classes, etc. Table IV gives the same summary for the second-grade classes.

We may now turn to the three questions which we raised near the beginning of our article. From Tables III and IV we observe that teachers agreed with test results in from 20 to 60 percent of the cases on which they passed judgment. Their disagreements, however, were much closer for dull children than they were for bright children. In the second grade about 53 percent and in the eighth grade about 63 percent of the teachers' estimates that pupils belonged in the lowest 20 percent were verified by the intelligence quotients. On the other hand, when the teachers judged that pupils belonged to the best 20 percent of their classes, they agreed with the intelligence quotients much less frequently—in but 22 percent of the cases in the second grade and in but 43 percent in the eighth grade.

It appears to be appreciably more difficult for teachers to judge the mentality of pupils in the second grade than in the eighth. Table IV when taken in comparison with Table III exhibits this fact so far as the teachers who participated in this investigation are concerned. In round numbers 20 percent more of the bright and 10 percent more of the dull children were correctly selected by the teachers in the eighth grades than in the second grades. In another investigation by the writer it was found that the teachers' estimates of the intelligence of 171 third-grade pupils correlated 0.43 with intelligence quotients while the estimates of intelligence of 89 seventh-grade pupils correlated 0.67 with the same measure. These correlations substantiate the conclusion to be drawn from this investigation, namely, that it becomes less difficult to estimate the intelligence of pupils as they become older.

To the extent that the results of intelligence tests may be accepted as valid and the judgments of these teachers may be regarded as typical, the following statements may be offered:

SUMMARY

1. Teachers can select 20 to 40 percent of the bright pupils in their grade and from 50 to 60 percent of the dull.

2. It is more difficult to select the bright pupils than it is to select the dull.

3. It becomes less difficult to select the bright and the dull pupils as they become older.

4. Selection of pupils for special classes by teachers' estimates is a very unreliable and unsatisfactory method of procedure and should not be used if it is possible to use an intelligence test. In the selection of pupils to enter classes for gifted children, teachers' estimates of intelligence are of practically no value.

5. From this study of teachers' ability to select the intelligent and the unintelligent children, inferences may be drawn as to their ability to select pupils according to other character traits.

THE MENTAL AGE OF ADULTS

EDWARD A. LINCOLN
Harvard University

The results of mental testing in the United States army during the late war left one vexing problem for the psychologist. The data, taken at their face value, indicate that the mental age of adults is not sixteen, as assumed by the Stanford Revision of the Binet test, but only about thirteen and a half. This point has become somewhat controversial, and has been recently attacked by Doctor Terman, who states, in the Journal of Educational Psychology (vol. XII, no. 6, p. 330), several reasons which, he believes, invalidate the inferences which may be drawn from the Army results. It is the purpose of this paper to inquire into the validity of Doctor Terman's objections to the inferences from the army results, and also to seek further light on the whole problem.

Among the objections which Doctor Terman offers, one of the most potent, if it can be shown to hold, is that "the 653 enlisted men[1] for whom the average mental age of 13.4 was found cannot be taken as representative of the entire draft army." The assumption is that these men were of inferior intelligence as compared with the draft army as a whole since, "A disproportionate number of them were from southern and semi-southern camps where, according to all the results of Army testing, average intelligence is lower than in northern and western states."

This assumption is not supported by the facts, which are shown in Table I. Here we find that the 653 group, when compared with the white draft as a whole, has a distinctly higher percent of A and B men, regardless of the examination upon which they are graded. The Alpha and Beta examinations also give the 653 group a larger portion of C+ men. If the sums are taken of the percents of men making grades above the average, that is, all A, B, and C+ men, the comparison again favors the 653 group. Table I also shows that in regard to the grades below the average there is little difference.

---

[1] In order to obtain data regarding the comparative scores of the several examinations used in the Army, requests were sent out from the Surgeon General to the psychological examiners in 12 camps that 100 in each camp, relatively unselected, be given Alpha, Beta, and the Stanford-Binet. These examinations were given for the most part in August, 1918, and of 748 men examined, 653 were born in English speaking countries. For full report see Yerkes, R. M., *ed., Psychological examining in the United States Army.* (Memoirs of the National Academy of Sciences, V. 15, 1921) p. 384 *ff.* Cited as "Army report" in this article.

TABLE I. PERCENTS OF WHITE DRAFT[a] AND OF 653 GROUP MAKING EACH GRADE

	Grades								
	D−, E	D	C−	C	C+	B	A	Above C	Below C
White Draft (94,004 cases)	7.1	17.0	23.8	25.0	15.2	8.0	4.1	27.3	47.9
653 Group									
Alpha.	18.1	8.4	19.3	21.3	18.2	8.9	5.8	32.9	45.8
Beta.	6.3	21.0	21.8	20.8	15.9	9.2	5.1	30.2	49.1
Stanford-Binet. . . .	8.0	15.6	23.0	22.4	13.2	11.9	6.0	31.1	46.6

[a] Figures for white draft from several tables in Army report, for example, Table 419 on page 865. Those for the 658 group are from Table 57, page 898.

Comparisons may also be made on the basis of the combined scale scores for the two groups, which may be found on Tables 332 and 333 on pages 789 and 790 of the report. In spite of the fact that the 653 group has no scores on the combined scale above 21.9 while the White Draft has 4.67 percent of its scores above this point, the smaller group has a slightly higher median; 13.5 as compared with 13.46 for the White Draft. This difference persists when translation is made into mental age; the median for the 653 is 13.25 years, as against 13.15 for the group from the principal sampling. It may be fairly argued, of course, that these differences are so small that they are not significant. However, this evidence, though slight, is cumulative, and with the facts in the previous paragraph it seems to indicate unequivocally that we cannot assume that the 653 group was inferior in general intelligence to the draft army as a whole.

A further objection which has been raised to the use of the Army results is the assumption that "the draft army itself was certainly not representative of the male population between the ages of 21 and 31," largely because it is believed that deferred classification and exemptions were, for the most part, granted to the men of superior intelligence. Data are not available for the proof or disproof of this assumption, but it is possible to point out some rather fundamental considerations which Doctor Terman omitted to take into account when he made it.

The Army report, on page 830, presents figures concerning the exemptions and deferred classifications which are gathered for convenience into Table II.

TABLE II. PERCENTS OF EXEMPTION AND DEFERRED CLASSIFICATION FOR VARIOUS REASONS

Dependency	51.1
Alien allegiance	13.16
In service	8.9
Disability	7.75[a]
Occupation	6.8
Religion	0.5

[a] The report says these men were "totally and permanently physically, mentally, or morally unfit..." This figure does not agree with the report of the Statistics Branch of the General Staff which states that from 15 to 50 percent were rejected for physical reasons. (p. 20-21).

Let us analyze as well as we may each of these groups, and find, if we can, any reasons for believing them to be composed of individuals who vary to any great extent on an intelligence basis from the population as a whole.

Over one-half the exemptions and deferred classifications were for dependents. Probably most of them were granted to men who were married and were supporting their wives and children. Is this group superior in intelligence to the population as a whole? It scarcely seems that we can make this assumption, and it would be hard indeed to prove it. Such evidence as there is seems to point the other way. The more intelligent man stays longer at his education, and is thus kept from early marriage. Moreover, he usually wishes a permanent place on a relatively high level before he takes the responsibilities of family life. The intelligent man is also more often successful in his work, and many of the men of this class were in such a position that they could not consistently claim exemption on the ground that they needed to continue their work in civil life in order to support their dependents. On the other hand, the less intelligent man, without ambition or foresight, marries early and has a large family. Thus it seems probable that the exemptions on this ground did not keep the intelligent man out of the draft army, but rather pulled him in, since it was more likely that he was unmarried, or if married, that his family was independent of his continued occupation in civil life.

Exemptions in the second largest group were for alien allegiance, that is, for those men who were not citizens of the United States. Are these men above the average intellectually? The Army report (part III, chapter 6, page 693) gives the results of the examination of some 12,000 foreign-born men who were included in the draft army because they had either become citizens or waived exemption. They were, as a whole, distinctly inferior to the rest of the draft army, as the figures in Table III will show. It is, of course, impossible to be sure that the foreign-born men who were drafted were on the same intelligence level as those who were not, but there seems no reason to believe that great differences existed. The only bit of evidence of which the writer is aware is that the conscientious objectors who refused to serve in the army because they claimed not to be citizens, because they were enemy aliens, or because they had enemy alien relatives fighting on the other side, were, as a group, of very low intelligence.[2] This would seem to indicate that the more intelligent foreigners came into the army, especially when

TABLE III. PERCENTAGE DISTRIBUTION OF LETTER GRADES MADE BY FOREIGN-BORN MEN IN THE DRAFT ARMY

	E, D−	D	C−	C	C+	B	A	NUMBER OF CASES
FOREIGN	14.8	30.8	16.5	26.6	7.3	2.9	1.1	12,407
WHOLE DRAFT	7.1	17.0	25.0	23.8	15.2	8.0	4.1	94,004

---

[2] Army report, p. 802, section 2.

considered with the fact that the men who were naturalized, and therefore liable to the draft, were in all probability superior to the average of the foreign-born. The same is likely to be true of those men who waived exemption. Thus in this field the weight of evidence seem to be that if the general population was not well represented by the draft army it was because the men in the army were of a higher, not a lower, intelligence level than the men who were not drafted.

Doctor Terman believes that the men exempted from the draft because they were already in the service formed a superior group because, "these volunteers included a disproportionate number of college graduates, college students, recent graduates of high schools, high-school students, and high-minded youths from the better classes generally." This may be true, but it cannot be accepted unless accompanied by adequate proof, first, that there *was* a disproportionate number of college, high-school, and high-minded youths in the group, and second, that those who volunteered were superior to the men of like education and high-mindedness who waited to be drafted. The declaration of war found some 200,000 men in the United States Army, approximately two-thirds of them regulars and the rest National Guardsmen, who had been called for service on the Mexican border.[3] Surely no one familiar with the type of man who joins these organizations will contend that, with all their virtues, these men constitute any part of "the intellectual cream of the country."[4] This group comprises about a third of those exempted because of prior enlistment. The others were men who had enlisted when the call for volunteers was issued, or sometime before the enlistments were closed. Can we declare that the men in this group who came, say, from Harvard University were more intelligent than the Harvard students and graduates who were later drafted? Decidedly not without some pretty clear evidence. There may be reason, indeed, to suppose that if any selective factor was acting it worked the other way, and that unsuccessful men, the adventurers, and the like, were the ones to rush into the army, while the superior men waited to be called. It is not without import, surely, that the percent of desertions from the army was smaller among the drafted men than among those who volunteered.[5]

---

[3] Report of statistics branch of the general staff, p. 16.

[4] In the preliminary tryout of examination *a* when the scores ranged from minus 188 to plus 226 the following results were obtained at Syracuse, Regulars —73; at Indianapolis, Regulars and National Guard —47; at Nashville, National Guard —35. These scores correspond to 115, 121, and 153 on the later examination *a* of 414 points, on which the median score of the enlisted men in the 4 trial camps was 151. Mosquito fleet men at Brooklyn tested +88 (176), but their inclusion in one group with the army men leaves the score —26 (162). Army report p. 315, Table 1, and p. 491, Table 145.

The exemptions for industrial and occupational reasons form a rather small group, but it is worth while to consider them as we have done the others. From the positions they held it seems quite legitimate to agree with Doctor Terman that they were superior individuals, but this argument does not take account of the fact that a man's position in his chosen field does not depend upon his mental ability alone. It may depend largely upon his father's ability, or upon the social, political, or financial connections of himself, or of any member of his family. There were charges during the operation of the selective service act that illegitimate exemptions were secured in this way. If we leave these considerations out, however, it is certainly clear that the attainment of the positions of leadership is due in a large measure, at least, to the length of time that a man has been working in his particular line. Men do not become "managers," "directors," "comptrollers" and the like until they have had a considerable degree of experience. Probably, then, the men who were exempted in this group were not more intelligent than their fellows, they were simply older.

Some evidence for this conclusion appears in the Army report in the chapter dealing with the intelligence ratings of occupational groups.[6] It is here stated that, "consistent significant differences in mental ability between apprentices, journeymen, and experts are not present." In only a few of the occupations do the experts score higher than the journeymen, and the journeymen higher than the apprentices, and many of the differences which do exist are small. Frequently the apprentices scored higher than the experts. This is exactly what would be expected from the considerations in the preceding paragraph. The exempted men were no more intelligent than the beginners; they were simply farther along in age and experience.

---

[5] Report of statistics branch of the general staff, p. 17.

[6] Army report, p. 831 ff.

The men exempted for physical, mental, and moral unfitness surely cannot be considered a part of the "intellectual cream of the country." They might possibly be considered as average, but even this is extremely doubtful. In the first place, this group contained those who were rejected on the basis of inferior mentality, and this fact alone would probably bring the group below the average. We know, too, the tendency of the mentally inferior to be physically inferior also, and undoubtedly some of the men rejected for physical reasons were of inferior mental calibre as well. For some groups of physical rejections we may be sure that the average mentality was low, since the Army report points out the mental inferiority of venereals and men infected with hookworm.[7] The sum of the evidence seems to show that exemptions in this group left the inferior men in the general population and put the superior men into the army.

The objection that the low scores were caused by unfavorable conditions of examination is refuted, in part at least, by the finding that recently inoculated men did as well as men whose inoculation was taken at a time when it was impossible for it to function at the time of the examination.[8]

Further arguments by Doctor Terman deal with the form of the Stanford-Binet which was given in the army, and the training of the examiners. These he presents without evidence. The abbreviated scale, he believes, gives lower mental ages for adults than the complete scale, and untrained examiners are likely to be severe in their scoring. On the first point it should be considered that the omitted tests were for the most part those which are linguistic in their nature. They were omitted because it was supposed that many of the men who would be tested by the individual methods could not do themselves justice on tests of this nature. Further, the tests remaining were, in many cases, those in which the men were able to make better records because of experience rather than mentality. The making-change test is a good example in this connection. It is a nine-year test in the Stanford revision, but it was passed by a large percent of the adults who tested only eight.[9] There is at least a possibility, then, that adults test higher rather than lower on the abbreviated scale, and it is almost certain that such would have been the case with most of the men in the army who could not be marked fairly on the basis of Alpha.[10]

---

[7] *Ibid.*, pp. 810-11.

[8] *Ibid.*, p. 809.

In connection with the second point it must be remembered that when the short army examination was made there were omitted many of the tests in which the examiner has to use his own judgment in scoring the response. For example, the omitted items included all the vocabulary tests, the meaning of abstract words, and the differences between abstract words. The fact of general severity in marking by examiners of small experience must also be called into question. It has never been noted in the training of students at the Psycho-Educational Clinic of the Harvard Graduate School of Education. We have found severity and leniency in marking to be largely a matter of the individual examiner.

The most serious objection to Doctor Terman's arguments, however, arises not so much from these arguments themselves, but rather from the fact that he failed entirely to take account of the striking evidence which arises from another source. He says nothing whatever concerning the fact that in many cases where the Alpha examination has been given to school children the thirteen-year-olds have equalled or surpassed the performance of the army men. The results of these testings in schools are especially significant because they are open to none of the objections which are raised in connection with the Stanford revision, and we need not even take account of the one unquestionably valid point raised by Doctor Terman, namely, that the revision may be too difficult in the upper ranges.

---

[9] *Ibid.*, p. 408. Also, Lincoln, E. A., and Cowdery, K. M., "An abbreviated mental age scale for adults," *School and Society*, 10: 708, December 13, 1919.

[10] The Army report on p. 399 discusses the reliability of the four-test scale, and seems to justify its use.

This superiority of thirteen-year-olds over army men appeared when the Examination *a*, the original group examination used in the army, was given to groups of California school children. It is to be remembered that the scores of all children foreign-born or of foreign-born parentage were eliminated from the distribution, but to offset this, in some measure at least, is the fact that the scores for the enlisted men are for only those who were considered literate. Care must also be taken not to fall into the error of considering the enlisted men as a group which scored low partly because it included no officers, since many of the men in this group were afterwards sent to officers' training schools and commissioned.[11]

In the Army report the following results of Examination *a* are recorded:

Median score 399 12-year-old children, grades 3–9	155[a]
" " 389 13-year-old " " "	182[a]
" " 274 14-year-old " " "	207[a]
" " 56,140 enlisted men, literate, from Camps Devens, Dix, Lee, and Taylor	161[b]

[a] Army report, p. 538, Table 186.
[b] Army report, p. 491, Table 145.

These figures seem to indicate that, even after considerable allowance for all sorts of factors, the army median cannot be considered to equal or exceed that of the thirteen-year-old children.

There are not in the Army report so far as the writer is aware, any results of the scores made on Alpha by school children. This examination, however, has been used in a number of schools since its publication, and the results have given further evidence that the mental age of adults is on a lower level than we had supposed before the army tests gave us a hint on the situation. The median Alpha score for 67,254 literate white men is 60.5,[12] and the median for 51,620 native-born whites is 58.9.[13] Doctor Doll examined with Alpha 514 children from the eight grades of the Franklin School in Trenton, N. J., and found as follows:[14]

---

[11] Doctor Terman falls into this error when he states that all officers were excluded from the 653 group.

MEDIAN AGE......	9.5	10.5	11.5	12.5	13.5	14.5	15.5
MEDIAN SCORE....	26	37	61	81	85	80	88
NUMBER CASES....	29	38	63	96	139	75	22

In this school the eleven-year-olds equal the army score, and again it seems that even with many allowances the army would not average higher than 13. It must be remembered in connection with the above figures that the medians for the upper ages are undoubtedly low, because a number of the more intelligent twelve- and thirteen-year-olds have gone on into the high school, while most of the fifteen-year-olds and many of the fourteen-year-olds left in the grades are mentally inferior.

A group of rural New York children selected from the fourth grade through high school and ranging in age from 8 to 21 gave the following scores:[15]

AGE...........	8	9	10	11	12	13	14	15	16
ALPHA SCORE...	35	35	41	46	56	81	86	95	99

This gives us a more accurate idea of the scores for the upper ages, and again indicates that the average soldier was certainly not much above 13 years old.

If we turn to high-school results, we find the following, reported by Sylvester and Dobbs.[16]

CLASS	MADISON		ROCKFORD		SIOUX CITY	
	Median Score	Average Age	Median Score	Average Age	Median Score	Average Age
Senior.........	121.2	17.3	123.8	17.5	137.6	....
Junior........	122.4	16.6	119.7	16.4	130.2	16.6
Sophomore.....	109.7	15.8	107.0	15.3	122.1	16.0
Freshman......	96.0	14.8	94.0	14.5	107.9	14.8

---

[12] Army report, Table 183, p. 666. This table contains a larger number of men than any other I can find.

[13] *Ibid.*, Table 309, p. 767.

[14] Doll, E. A., "The Average Mental Age of Adults," *Journal of Applied Psychology*, 3:317-28, December, 1919.

[15] Lufkin, H. M., "Report of the use of the Army Alpha test in rural schools," *School and Society*, 13: 27-30. January 1, 1921.

[16] Madsen, I. N., and Sylvester, R. H., "High school students' intelligence ratings according to the Army Alpha test," *School and Society*, 10: 407-10, October 4, 1919.

We must not forget in connection with these figures that the high-school pupils are a select group, but it does not seem that selection alone can account for the fact that the scores of the fourteen-year-olds are 50 percent and those of the sixteen-year-olds more than 100 percent above the average score of the army.

A valuable contribution in this field has been made by Kohs and Proctor.[17] They report a set of Alpha norms based upon the examination of several thousand California school children. Of this group they say, "The children tested were drawn from all types of communities, and represented every variety of social status." That part of the table which presents scores for twelve- and thirteen-year-olds is here reproduced:

ALPHA SCORE	MENTAL AGE NORMS	
	Kohs and Proctor	Army
45	12–0	13–0
50	12–4	13–3
55	12–8	13–6
60	13–0	14–0
65	13–4	14–3
70	13–8	14–6

50.1 Median of 653 group
58.9 " 51,620 native whites
60.5 " 67,254 literate whites

These results indicate that the army mental age norms were somewhat high, and that the average man may have been slightly *under* thirteen mentally.

So the evidence piles up. The writer, however, does not wish to convey, through this article, his conviction that we should, without further investigation and analysis, adopt 13 or 14 instead of 16 as the average mental age for adults. His purpose is to set forth the fallacies in the arguments now generally advanced to invalidate the army results, in order that the question may be kept open and studied scientifically. If we are to use the results already in hand, the crux of the matter lies in the answer to the question: To what extent was the army typical of the population at large? An analysis of some of the evidence has been presented in this article. Much further light can be obtained by the study of the psychological records now on file at the National Research Council. The study might well begin with the examination of the records of the men who were only twenty-one or twenty-two years old at the time of their entrance to the army. In this group the exemption for marriage and occupation would probably be low, and we would be rid of one troublesome factor. The examining of school children should continue, with care that the examinations are given far enough down in the grades, and far enough up into the high school to include all the children of ages 12, 13, 14, and 15, thus avoiding any factors of selection.

---

[17] Proctor, W. M., "*Psychological tests and guidance of high-school pupils,*" Journal of Educational Research Monograph No. 1, June 1921.

ATTEMPTS AT TEST VALIDATION

RAYMOND FRANZEN
Director of Research, Des Moines Public Schools

Measurement in education has passed its infancy. We no longer need to plead for recognition of the cardinal importance of finding exact indices of the quality which we are describing. Our attention has been diverted from the question of whether it shall be done, to the question of how it shall best be done. The increasing popularity of tests promises an early verdict from the school men of America regarding the practical consequences and immediate advantages of such measurements. This verdict will be transmitted to us in terms of the tests that are used, not in terms of the tests that could have been used. It is, therefore, of utmost importance that the tests be more carefully constructed and fairly evaluated.

Administration is not scientific in its choice of tests. Geographical preferences, advertisement, and the authority and prestige of the author are prominent elements in evaluation. Administrative exigencies, such as the price of a test, the time it takes to give it, and the convenience of its scoring facilities, will provide a sale for tests whether they have other virtues or not. Such considerations should, of course, be entirely secondary, since price, time, and convenience can only be of value if the test renders the service that is needed; and exact statistical virtue must form the basis of sound judgment. Cost should be considered only *in relation* to value.

Chief among the fundamental virtues which determine the value of a test are, first, that the test shall measure what it purports to measure, and second, that it shall be reliable. It is very important that a test have objectivity, that it be scaled, that it have usable standards, and that it exhibit certain other secondary characteristics; but if it is to be a test at all, it is essential that it tell the story which it pretends to tell, and that it tell this same story upon similar, successive occasions.

An attempt is here made to determine in respect to entering high-school pupils, the degree to which our so-called intelligence tests measure what they purport to measure. In this connection I may say that I believe the effect on measurement would be salutary if in naming our tests we used less readily such a general term as "intelligence." We might well substitute for it some expression descriptive of a feature of the criterion against which, in order to prove its usefulness, the test is checked. Thus a kindergarten test which attempts to predict future success in first grade, and whose value is judged upon a basis of the correlation of the test given in the kindergarten with future success in first grade, should be called "A Kindergarten Test of First Grade Readiness," rather than "An Intelligence Test."

It is very difficult to decide what elements of school life to use in validating our "intelligence tests." It is quite within the range of possibilities that the various authors of tests differ radically in their conception of what minutiae of behavior combine to form the power of adaptation which by their tests they are attempting to measure. It is, therefore, very difficult to choose criteria which will tell us how nearly the tests do the job. A more intricate and refined evaluation of any one test is possible in various ways, but a wholesale determination of relative value means the choice of some easily obtainable criteria. An experimental criterion including a careful observation and accurate clinical description of children is impracticable. The use of an individual Binet test is not essential for this purpose since the average of a dozen group tests would in itself probably be as good a measure of the intelligence of children as the results of a Binet test. The unreliability of marks given by teachers, and of the judgments of the intelligence of children by teachers, makes these unsatisfactory media for comparison. Age, used negatively, though it is an indication of how long it has taken the children to get as far as they are, is insufficient, because children are promoted for reasons other than readiness to go on. The use of the totals of all the tests is to some extent reasoning in a circle, though nearly everyone would grant that the average of twelve or fifteen tests of power of "adaptation to new material" is a fairly accurate indication of what is usually meant by intelligence. We should be willing at least to admit that a test which correlates positively and high with the total of a number of tests is a better instrument than one which correlates either postively and low or negatively—though this might mean that the test in question measured something different. However, if we do not know what this "something different" is, this is an almost useless reservation. The intercorrelations and the partial correlations of a test with the total, irrespective of the others, is another factor which contributes to evaluation, though the use of partials would logically follow the determination of a very acceptable criterion so that we might then select the best team of tests to do the work which we have defined in this acceptable criterion.

The use of these elements as criteria will establish in a rough manner the efficacy of the tests. In this investigation we have examined fourteen intelligence tests with respect to such criteria. Table I indicates certain correlation data. The spread of the data is not great; largely because the material concerns but one high-school class. The average age of the fifty-seven pupils who took all these tests was 168.82 months, and the standard deviation of their ages was 10.73 months. This point —namely, the variation in age—is important. We do not so much wish to find how well a test distinguishes a normal ten-year-old from a normal thirteen-year-old as to find out how well it will distinguish a stupid thirteen-year-old from a bright thirteen-year-old. We need an instrument which will distinguish neural quality when neural change due to age is constant.

Some of the criteria above mentioned have been used here in spite of their insufficiency in order to make crude distinctions between our better and our worse media of measurement of intelligence, leaving for a more detailed and intimate investigation the discrimination between the better ones, and the determination of what it is that the others test. These data are at best a mere indication of even this crude distinction. As I

TABLE I.[c] EACH TEST CORRELATED WITH THE TOTAL AND WITH EACH OF THE OTHER TESTS

	TOTALS OF ALL TESTS EXCEPT THE WYLIE	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. NATIONAL-A......	.93														
2. TERMAN-A.......	.92	.85													
3. OTIS............	.92	.93	.86												
4. HAGGERTY........	.91	.83	.85	.78											
5. ILLINOIS GENERAL	.90	.85	.83	.90	.81										
6. PRESSEY-SURVEY..	.87	.81	.82	.37	.76	.78									
7. MENTIMETER......	.81	.72	.89	.73	.82	.67	.69								
8. THORNDIKE READING[a]........	.81	.77	.72	.83	.77	.84	.63	.44							
9. NATIONAL-B......	.80	.76	.76	.73	.73	.80	.61	.60	.51						
10. DEARBORN-1......	.74	.64	.68	.60	.69	.52	.65	.64	.37	.48					
11. DEARBORN-2......	.72	.67	.70	.55	.66	.55	.64	.63	.39	.59	.60				
12. PRESSEY CROSS-OUTS[a]...........	.70	.69	.58	.74	.66	.66	.50	.58	.48	.63	.56	.43			
13. WYLIE[b]..........	.64	.55	.67	.65	.64	.60	.59	.54	.63	.48	.57	.32	.37		
14. MYERS...........	.63	.53	.57	.56	.49	.49	.54	.57	.25	.39	.54	.36	.35	.34	
STANDARD DEVIATION.	15.76	16.65	28.66	24.21	18.43	20.25	15.46	11.53	.45	14.25	12.28	12.47	11.02	18.35	15.17
MEAN...............	143.60	143.70	113.53	136.04	122.61	106.95	139.91	103.28	7.24	146.09	64.46	64.84	73.67	85.00	57.30
COEFFICIENT OF VARIABILITY......	0.11	0.11	0.25	0.18	0.15	0.19	0.11	0.11	0.06	0.10	0.19	0.19	0.15	0.22	0.26

[a] Thorndike Reading and Pressey Cross-outs are 20-minute tests.

[b] Wylie A, B and C is a 15-minute test.

[c] There were 200 cases to begin with, but N of these correlations is 57—being the number that took all the tests.

have already stated, there are only fifty-seven cases involved, but the number of correlations used, the interpretations of which supplement each other, makes this a less serious objection. We started with a whole freshman class in one high school, about two hundred, but before the administration of all fourteen tests was completed the number taking all examinations was reduced to fifty-seven.

Table I gives the correlation of each test[1] with the total (divided by ten for ease of statistical computation), and the intercorrelations of all the tests. The total was obtained by adding the scores of all of the tests except the Wylie. Scores on this test had not yet been tabulated when the total was obtained. Each test was roughly equated with the others in variability before the addition. Table I includes for each test the standard deviation, mean, and coefficient of variability (standard deviation divided by mean).

Table II gives the correlations of each test with the total, and the intercorrelations of all tests, with reading ability, as measured by Thorndike Alpha-2, rendered constant.

Table III gives the average of the correlations of each test with the other thirteen. This table is a rough check on the first column of Table I.

Table V gives the intercorrelations of the elements used as marks received the first semester in high school. Column 4 of the same table gives the correlations with the average of teachers' judgments of the child's power to adapt himself to his environment. These judgments were rendered after at least half a semester's acquaintance with the children. Column 6 of Table IV gives the correlations with age at high school entrance, or the length of time it took the children to get to high school. Naturally these correlations are negative and the higher the negative "*r*" is, the better the test is in that respect.

Table V gives the intercorrelations of the elements used as criteria. As to the method of obtaining the teachers' judg-

---

[1] **The Dearborn examination for higher grades was divided into two parts since it was twice as long as an ordinary test. The first half was called Dearborn-1 and the second half was called Dearborn-2.**

TABLE II. INTERCORRELATIONS OF TESTS WITH THORNDIKE READING RENDERED CONSTANT (N = 57)

	TOTALS	1	2	3	4	5	6	7	8	9	10	11	12
1. MENTIMETER	.87												
2. TERMAN	.85	.92											
3. NATIONAL–A	.84	.66	.68										
4. DEARBORN–1	.81	.57	.64	.60									
5. PRESSEY–SURVEY	.80	.59	.70	.65	.58								
6. HAGGERTY	.78	.83	.68	.59	.68	.55							
7. OTIS	.78	.72	.67	.81	.56	—.35	.39						
8. NATIONAL–B	.78	.49	.66	.67	.36	.44	.62	.65					
9. MYERS	.77	.53	.58	.55	.50	.51	.48	.65	.31				
10. DEARBORN–2	.75	.55	.67	.63	.53	.55	.61	.44	.49	.29			
11. ILLINOIS	.71	.61	.62	.57	.42	.60	.46	.67	.80	.54	.44		
12. CROSS–OUTS	.61	.47	.38	.57	.46	.29	.52	.69	.51	.27	.30	.54	
13. WYLIE	.24	.38	.42	.12	.47	.32	.31	.30	.24	.24	.10	.17	.10

ments regarding the ability of children to adapt themselves to the environment the following instructions[2] to the teachers were used:

> "I desire a special report upon each of the pupils whose names are upon the list read in teachers' meeting. This report is to be made up by each teacher who has had the pupil in class for more than one-half semester. Use the slips of paper furnished you, making out a separate slip for each pupil. Write his name upon the slip and under it your score according to the following directions:
>
> "Using a scale from zero to ten, indicate your estimate of this pupil's native ability, that is, his power to adapt himself to his environment. This may be interpreted to mean his natural ability to solve a new problem without assistance. Try to eliminate from your estimate all acquired characteristics, in other words, pay no attention to such questions as to whether he is lazy, is poorly prepared for his work, is timid, has bad habits of study, etc. A score of zero means that the pupil has no power of adaptation to the circumstances which surround him in your class.
>
> A score of ten (10) would mean that this power was as nearly perfect as you would consider possible to a child of his age, or at least just as good as that of the brightest child you have ever known."

Some of the tests show a very good rank with respect to each of these various ways of viewing their value. The Terman for instance, is rather high in each of the lists. The Illinois is high in each except one, Table II, which means that it is a very good medium of prediction, but that it measures reading as well as intelligence (as far as the Thorndike test measures reading). The Thorndike Reading, though it is only a twenty-minute test, stands high in each of the tables, except Table II, in which it is of course, not included. The Otis, the Haggerty, and the National tests give a fairly good account of themselves. The Mentimeter has the highest correlation with the totals of all other tests, when reading is rendered constant. This would justify us in saying that if we wish to use a non-verbal test, we should use the Mentimeter. The Wylie stands very high in the tables, in consideration of the fact that it represents only fifteen minutes of testing. Its high negative correlation with age indicates that it measures position within an age group better than growth itself. The presence of this desirable feature in the Wylie is verified by other data showing a wide spread within age groups. The standing in the tables of the Myers and the two Dearborn tests indicates the comparative inferiority of these tests in terms of criteria as here defined. The two Pressey tests vary too much in the different tables for us to risk a judgment.

---

[2] These instructions were prepared by Mr. A. W. Merrill, the principal.

Table II can be usefully employed to give indications of the degree to which the prediction of a test is valuable, apart from the prediction which is rendered by reading. It shows association of the remnants when agreements in terms of reading are "ironed out." We are most interested in that kind of prediction, since school marks and teachers' judgments of intelligence are based largely on ability to read, and we want to know those who are capable but who cannot read. The Thorndike Reading test and the Illinois General Intelligence test correlate highest with teachers' judgments because they both involve reading to such a great extent. The Illinois General correlates so high with the reading test that we may say that it is a measure of reading.

The intercorrelations as they stand in Table I and as they stand in Table II, with the reading partialled out, are valuable to determine which tests to use together. Unless it is merely a more reliable result which we are seeking, we should not use tests together which are highly correlated. If the correlations presented in Table I are valid, Otis and National-A should not be used together. Terman, Thorndike Reading, and Cross-outs or Wylie would be a good team, since each is of some value according to these criteria and their intercorrelations are low.

It is not out of place here, to add that the size of the correlations with these criterion elements is no reflection on the tests, but rather on the elements themselves. Even if we were to admit that the school marks are a desired goal and even if these marks were reliable enough to yield a high correlation, we should still be rendering a better prediction of this kind of success by our best intelligence tests than expert corn judges render of the yield of corn when they judge it at our Iowa fairs.

TABLE III.—AVERAGE OF CORRELATION WITH THE OTHER 13 TESTS

TERMAN	.75
NATIONAL–A	.74
HAGGERTY	.73
ILLINOIS	.72
OTIS	.71
MENTIMETER	.66
PRESSEY-SURVEY	.65
NATIONAL–B	.62
THORNDIKE READING	.59
DEARBORN–1	.58
CROSS–OUTS	.56
DEARBORN–2	.55
WYLIE	.53
MYERS	.46

TABLE IV.—CORRELATION OF EACH TEST WITH CERTAIN CRITERIA

TEST	WITH AVERAGE OF H. S. MARKS THE FIRST SEMESTER		WITH AVERAGE OF TEACHERS' JUDGMENTS OF "POWER TO ADAPT"		WITH AGE	
	r	Rank	*r*	Rank	*r*	Rank
1	*2*	*3*	*4*	*5*	*6*	*7*
THORNDIKE READING	.51	1	.67	1	–.42	4
ILLINOIS	.42	2	.62	2	–.47	2
NATIONAL–B	.41	3	.55	3	–.33	7
OTIS	.31	4	.52	4	–.45	3
CROSS–OUTS	.30	6	.41	13	–.29	10
TERMAN	.30	6	.51	5.5	–.39	5
WYLIE	.30	6	.48	8.5	–.49	1
HAGGERTY	.29	9	.50	7	–.32	8
DEARBORN–1	.29	9	.48	8.5	–.24	12
NATIONAL–A	.29	9	.47	10	–.30	9
MENTIMETER	.26	11	.45	11	–.27	11
DEARBORN–2	.24	12	.42	12	–.12	14
PRESSEY–SURVEY	.22	13	.51	5.5	–.38	6
MYERS	.12	14	.28	14	–.14	13

TABLE V.—INTERCORRELATIONS OF THE ELEMENTS USED AS CRITERIA. AGE IS USED NEGATIVELY

	MARKS	JUDGMENTS	TOTALS
JUDGMENTS	.82		
TOTALS	.34	.58	
AGE	—.37	—.48	—.40

The correlations of four judges' scores of ears of corn with the yield of those ears were all about 0.20.[3]

The intercorrelations of the totals with the other elements of the criteria given in Table V indicate the limitations of these other elements. For example, since the average of teachers' judgments correlates only 0.58 with the sum of 13 tests, there is undoubtedly something wrong with the judgments. Part of their insufficiency as validation criteria is indicated by their high correlation with the pupils' marks. The marks are of June, 1921. The judgments are of December, 1921. They correlate 0.82. And still the judgments correlate only 0.58 with the totals of 13 tests. This means that the teachers' estimates of intelligence as defined in the instructions were very much more like their estimates of the product attained in their courses than they were like actual manipulation of new material by those children. Besides this, as the teachers' judgments are one element in these correlations and their reliability is very low, we cannot expect very high *r*'s. A correlation coefficient is limited in size by the reliabilities of its elements. The correlation of any test with others cannot go higher than the square root of its correlation with itself.

These tables attempt to answer three questions in a limited and crude manner.

(1) Which tests are most likely to be good for all-around use? The answer obtained from these data is Terman, Alpha-2 Reading, Illinois, the Nationals, Otis, Haggerty, and if a cheap and short test is wanted, the Wylie Opposites.

(2) Which tests are best to use if intelligence, irrespective of reading, is the objective?
The answer of these data is Mentimeter, Terman and National-A.

(3) Which tests should be used together and which test the same qualities?
The answer is to be found in Tables I and II.

---

[3] These data provided by H. A. Wallace of Des Moines.

More extensive and more intensive investigations of this nature will yield us fairly accurate determinations of the relative value of intelligence tests. In addition we must carefully demonstrate the quality of the quantitative messages that our product tests transmit. Here our methods are more clearly outlined. It is even possible in some cases to decide by analysis that a test measures some definite quality though the limits of such a method of validation are very restricted. Arithmetic tests do test arithmetic on the face of it, though it is instructive to compare Woody, Woody-McCall, and Courtis; and then, bearing in mind their differences, to hear teachers and principals talk of arithmetic scores, not "Courtis" scores or "Woody" scores.

Three of our well-known reading tests correlated as follows in one class of 40 children:[4]

	Thorndike-McCall	Monroe
Monroe	.26	....
Burgess	.19	.48

Although much more work with these tests is necessary to determine their true correlation, it is obvious from the data at hand that we need to investigate our reading tests. In a recent number of the Journal of Educational Research Courtis says:

> "The child who makes a high score in my test and a low score in McCall's may need to be given a series of life experiences rather than a series of reading experiences. To call McCall's a reading test is to imply that the child's deficiency can be made up by practice in reading. This is not true."[5]

I do not know whether this criticism is meant to extend to the Thorndike Alpha-2 or not. The type of the test is exactly the same. But I do know that a child's deficiency in the Alpha-2 can be made up by practice in reading. In an investigation at Garden City, New York, the correlations and differences between Intelligence Quotients and Reading Quotients

---

[4] These are correlations obtained by Miss Minnie Thurber of the Public Schools of Los Angeles, California.

[5] Courtis, S. A. "The analysis of reading ability." *Journal of Educational Research,* 4: 292, November, 1921.

(Thorndike Reading Age, divided by Chronological Age) all showed increasing correspondence of Binet scores and Alpha-2 scores after reclassification and training in *reading*, not life experiences.

This is one important way of validating a test. Does the correlation of the test with mental ability approach unity when training in the subject is given? If so, the product test and the mental test used are adequate. Woody-McCall Arithmetic, Thorndike Reading, and Stanford Revision Binet do test what they purport to test by this standard.

May Ayres Burgess in the 20th Yearbook of the National Society for the Study of Education says:

> "Since TB is the usual abbreviation for tuberculosis, print the two initials closely contiguous to this celebrated emblem of the American Tuberculosis Association. In doing so, be careful to select such locations as will result in the separation of the initials by the intervention of the emblem between them."
>
> Third-grade children, Italians, on the lower east-side of New York City, who were counted by their teacher as exceptionally poor in reading, and were not themselves able to use English fluently, read that paragraph, re-read it, and finally with shining eyes grasped their pencils, and printed T on one side and B on the other side, closely crowded against the central cross. They could not have defined the words, but they could discover what the paragraph meant if given long enough time to puzzle it out. The same result occurred over and over, with other forms of test material, even such as, under the picture of the veiled lady, requested the children to "Perpetrate an incision in the facial vesture immediately anterior to the buccal orifice through which the fastidious female may procure aliment."[*]

But when such a picture as this is put before third-grade children without any reading matter and the examiner merely says, "Do something to the picture," many of the children do just this.

If Picture Supplement-1 is given orally (that is, if only the pictures are put before the children and the instructions are given orally), and if Picture Supplement-2 is afterwards given according to instructions, except that plenty of time is allowed, the correspondence of the results is marked. Third-grade children fail both ways; sixth-grade children get very nearly perfect scores with little spread in both ways of giving. Then reading has not much to do with it. In the lower grades, it is a test for following directions, and in the upper grades it is a test of speed of doing what one is asked to do, orally or in print.

---

[*] Burgess, May Ayres. "Controlling factors in the measurement of silent reading." *Twentieth Yearbook of National Society for Study of Education*, Part II: pp. 29-30.

Miss Kate Kelley, the primary supervisor of Des Moines, and I are constructing a reading test for I-B children, in which the first step in the process of test building is the administration of each proposed element orally, and the inclusion of only those that *practically all children get correct when given orally.* Then, all the difficulty is reading difficulty. It is impossible to guess which elements will fall low when given orally. Who could predict that when orally presented about 90 percent of first-grade children would say that "The sun sets in the morning" is true, or that "Boys and girls never sing" is true? Still this is a reliable determination, whereas seemingly harder sentences are correctly marked by about 90 percent of first-grade children.

When all sentences included are correctly marked by at least 90 percent of the children when given orally and from 50 to 100 percent when given in print, then it is fair to attribute the added difficulty to reading. This is a second possible method of validation.

Pressey[7] has suggested that we use the lack of overlapping of two distinct groups as a measure of validity. For example, when National-A is given to all children in high school, who received two or more of the lowest marks given, and to those who received two or more of the highest marks given, the overlapping is small. Only one or two cases out of 100 of the group marked high, fall below the average score of the group marked low, and only one or two out of 100 of the group marked low, fall above the average score of the group marked high. A study of these children demonstrates why. This is a third possible process of evaluation.

---

[7] Pressey, S. L. "Suggestions looking toward a fundamental revision of current statistical procedure, as applied to tests." *Psychological Review,* 27: 466-472. November, 1920.

Teachers' judgment may be used to better advantage than merely as a measure against which to correlate test marks. For example, as a partial validation of the Park-Franzen test for readiness to do first-grade work, we had each of the fifty kindergarten teachers at Des Moines judge this readiness of her children; and we then correlated the test results for each class against the judgments of the teacher of the class. Thus we obtained fifty correlation coefficients. Each of them, however, concerned a series of judgments by a single teacher. Then arose the question of the ability of these teachers to make correct judgments; and, as the best answer to the question, Miss Park ranked them according to her estimate of their ability in this respect. If the correlation coefficients between test results and teachers' judgments were highest for the classes having teachers of best judgment and lowest for the classes having teachers of poorest judgment—in other words if the coefficients showed a high correlation with the teachers' ability to judge—this fact would be a favorable indication for the test. In other words, if there is a direct relation between a teacher's ability to judge, according to her supervisor, and the size of the correlation of her judgments with the test, it is a favorable comment on the test.

The intercorrelations with other tests is always a valuable determination; and whenever a criterion built up by careful clinical observation and repeated individual examination is possible, this is the preferable technique.

We are facing an era of selection of tests. We want fewer tests for measuring achievement in a given trait and better construction of such tests.

Editorials

REACTION

In the world of force and energy, there is a law that to every action there is an equal and opposite reaction. Students of history have pointed out that a similar law seems to operate in the growth of men and institutions.

In both plants and animals and in man himself, growth is neither uniform nor continuous. Periods of rapid change are followed by relatively longer periods in which very little growth is noticeable. So it is with national life and with social readjustments. In the few short years of the world war more new ideas were spread broadcast among all peoples than in any previous period in recorded history. It is inevitable that the reaction of the world to these ideas should be correspondingly profound; and it will be quite normal if, when the social order shall have stabilized on the new level, there should follow a long period of peace and prosperity.

Reaction, however, has sometimes a more sinister meaning. It has happened in human history that changes which have cost much in effort and blood have been completely wiped out by the reaction which followed so that the old order was not only reinstated but reinforced, and the coming of progress correspondingly delayed. Not every revolution against a tyrannical government has resulted in greater freedom for the revolutionists. More than one king has been able to stage a "come back." So with many minor details of progress. More than once the first experimental trial of a new procedure has lead to a reaction which wiped out every vestige of the new idea for at least a generation.

It is one of the outstanding characteristics of the times in which we live that the rate of progress has been enormously increased, and at the same time the periods of reaction have been correspondingly reduced. Today change, progress is almost continuous. The most radical inventions, the most far-reaching changes hold the public mind only for a few short days. Their good points are quickly realized and they themselves are at once incorporated into our daily living. Such reaction as there is, is largely regret by the older generation for the passing of some phase of social custom made superfluous by the new, or objection to some new custom in which they do not share. Is a reaction against the telephone thinkable? Is it probable that the use of automobiles will some day be forbidden and the day of the horse-drawn vehicle restored? How long will the aeroplane and wireless remain mere marvels? Even now they are being put to commercial use, although such use involves a change of social relations and will result in social readjustments so profound that a few generations ago they could not have been brought about except by revolution and bloodshed.

The secret of the assimilative powers of the present generation is education. The invention of printing made universal education possible. Education in turn made possible direct appeal to individual thought and reason. Today publicity for each new invention, for each social change accomplishes in a few short days what formerly it required several generations to bring about. Where there is an efficient "selling campaign" which reaches all the individuals involved, there is no possibility of reaction.

For education itself this means that progress is dependent upon the organization and maintenance of effective forms of training for teachers in service and equally effective forms of letting parents know the reasons for the changes made. Changes in social life compel changes in subject-matter and in methods of teaching; and the education which the school provides is not permitted to lag far behind the education needed for the social life of the day. Society cannot wait for the slow processes of natural evolution to replace old teachers with the new. On the other hand, to make changes which are not understood by either the teaching force or the public is to invite reaction.

A single illustration will prove the point. Twelve years ago the first reliable instrument for the scientific measurement of educational products was invented. Ten years ago standard tests were used for the first time in a great city survey. Now tests and scales are to education what a ruler is to a carpenter, what a thermometer is to a physician. Their value has been established beyond all question. The contributions to education made by measurement have been many and far-reaching. The teacher who does not measure, who does not base her thinking upon the evidence obtained from scientific experimentation bears about the same relation to the one who does that the old country practitioner bears to the modern diagnostician. Today the teachers who are being graduated from our teacher-training institutions are not even conscious that they are measuring. Measurement has become an integral part of teaching and is accepted by the rising generation precisely as they accept the telephone and the automobile.

But what about the teachers in service and the general public? Rumblings of reaction are heard from time to time and progress is delayed. Fortunately, very fortunately, the teachers themselves have realized the situation and have crowded into summer schools and regular sessions of our universities and colleges to get for themselves the newer ideas and methods. Fortunately, the children carry word from school to home and serve to educate the parents. Fortunately, also, the children have accepted tests and realized their value. Thus it has come about that a great change has been successfully effected in education, and the danger of disastrous reaction avoided.

That school authorities can always hope to be equally fortunate by chance is improbable. Both the public and the profession must come to see that there can be no stable progress, no freedom from blighting reaction without deliberate, well-planned publicity, mutual understanding between the home and the school, and adequate provision in time and opportunity for the continued training of the teachers in service.

S. A. C.

Reviews and Abstracts

Stout, John Elbert. *The development of high-school curricula in the North Central States from 1860 to 1918.* (Supplementary Educational Monographs, Vol. III, No. 3.) Chicago: University of Chicago Press, 1921. 322 pp.

The author first reviews the earlier high-school movement of the period previous to 1860. He shows that the curricula of the high schools changed slowly during that period. The aim of the Boston High School was for general education rather than preparation for college, while that of other schools was twofold, thus determining the type of school that has now become prevalent in the United States. The curricula of this earlier period covered a wide range of subjects and were not uniform although the differences were not striking.

For convenience the following period of 40 years is divided into eight five-year periods. A series of tables is presented made up of a study of the curricula of an increasing number of schools for each period. Little variation is found in the content of these curricula. Toward the close of the period some courses in commercial subjects appeared, but not to a sufficient degree really to characterize a curriculum. The same was true to a still less extent of normal courses for teachers.

The dominant course took the form of preparation for college with certain parallel elective courses in which some one subject, as science, might be emphasized. But there was no uniformity practiced aside from college preparation. The arrangement of prescribed and elective courses made its appearance more recently. The non-preparatory course with specific aim came very slowly.

The order in which subjects appeared in courses was practically the same as that still observable with respect to the older, so called, academic courses. Up to the close of the nineteenth century the prevailing custom was to divide the year into three terms. The semester division is a development of the last twenty years.

Chapters VI, VII, and VIII are devoted respectively to a discussion of "Constants and Variables," "Time Devoted to Fields and Subjects," and "Influence of Size and Location of Schools."

Part II treats of conditions and changes in subject-matter. Especial attention is here given to English, Sciences, the Social Studies, and Foreign Language and Miscellaneous Subjects. To students of the present day these studies are very enlightening as indicating the courses of work and the trend of development in the fundamental fields of high-school instruction.

Part III deals with "More Recent Developments," in which the writer clearly analyzes the general character and significance of the relatively rapid change that has taken place during the last fifteen years.

On the whole this is a valuable contribution to historical material in secondary education and the work shows a thoroughness and insight that is perhaps too exceptional in the subject-matter of many educational publications of the day.

H. A. HOLLISTER

University of Illinois

THORNDIKE, EDWARD L. *The psychology of arithmetic.* New York: The Macmillan Company, 1922. xvi and 314 pp.

"Satisfyingness in results" is a chief factor in fixing bonds. What attachments one forms for his old friend "Psychology," and what pride in the present dynamic subject, when he reads Thorndike's "Psychology of Arithmetic!" Professor Quack and Mr. Charlatan may tell us (free?) how "psychology" will keep us young and beautiful, and how to hold our husband's love; but Thorndike for little or nothing lays bare the inner workings of the teaching and learning process that thousands of teachers may become more skilled and millions of children more efficient. The "medicine man" in psychology may obtain his knowledge and inspiration in a week's correspondence course, but twelve years ago the writer sat in Professor Thorndike's classes in the "Psychology of Elementary School Subjects" and heard this treatise on arithmetic "in the making." It has been "in the making" for many years since and the modesty of the author's claim is in significant contrast to the psychological bootleggers above, when in the preface it is stated: "So they (the lectures) are now published in spite of the author's desire to clarify and reinforce certain matters by further researches!"

Although not at present able to define in detail the formation of such bonds as are more generally stated as ability to read, spell, add, subtract, etc., "the psychology of school subjects begins where our common sense knowledge of these functions leaves off." In terms of problems the task, however, may and should include a study of the following topics:

1. What is the function—just what in detail is an understanding of decimal notation, the moral after-effects of literature, etc.?
2. How are degrees of ability or improvement in these functions measured? For example, what degree of spelling word difficulty or knowledge of a fraction should we attempt in grade IV?
3. What can be done toward securing better control of the particular situation—response connections?
   For example, what are the similarities or dissimilarities between the bonds formed in learning 7 and 4 are 11, 7 and 5 are 12, and 27 and 4 are 31, 27 and 5 are 32, etc.?
4. What sets of bonds produce most advantageously the desired changes in knowledge, skill, or powers? Should the pupil be taught that "0 and 0 are 0," "0 and 1 are 1," etc., or taught the habit of entering 0 to a column of zeros and neglecting the 0 in general in addition?

5. What degree of strength is most desirable in bonds at given stages of learning? Should certain denominate number tables be learned at their first appearance to be remembered a month, a year, or for the work at hand?
6. What is the best order to follow in fixing sets of bonds when change is desired? In early writing habits, should details of stroke precede or should children "write" words and sentences? Metric system bonds are usually found late in the course. Should they be formed early enough to facilitate knowledge of decimals?
7. What are the original tendencies and pre-school acquisitions upon which the school may base or counteract its connection forming?
8. What armament of satisfiers and annoyers, of positive and negative interests and motives, stands ready for use in the formation of uninteresting connections between black marks and meanings, etc. "School practice has tried, more or less at random, incentives and deterrents from quasi-physical pain to the most sentimental fondling, from sheer cajolery to philosophical argument, from appeals to assumed savage and primitive traits to appeals to interest in automobiles, flying machines, and wireless telegraphy. Can not psychology give some rules for guidance, or at least limit experimentation to its more hopeful fields?"
9. How do general laws of efficient learning apply in the case of each elementary school task? Self-competitive arithmetic drills have proved effective in interest and results. How may other arithmetical functions be made to conform?
10. How far are special abilities or difficulties inherited?

Such are the problems. The book discusses these with a wealth of concrete illustration in the field of arithmetic. The author has literally laid bare its beginnings and processes. At the close of the chapter on "The Nature of Arithmetical Abilities" he makes a significant comment: "What may be called a 'sociology' of arithmetic is very much needed to investigate this matter. For rare or difficult demands the elementary school should not prepare; there are too many other desirable abilities that it should improve. . . . It seems probable that the school training in arithmetic of the past has not given enough attention to perfecting the more elementary abilities. And we shall later find further evidence of this. On the other hand, the fact that people in general do not at present use a process may not mean that they ought not to use it. . . . When we have an adequate sociology of arithmetic, stating accurately first who should use each arithmetical ability and how often, we shall be able to define the task of the elementary school in this respect. For the present, we may proceed by common sense, guided by two limiting rules. The first is—'It is no more desirable for the elementary school to teach all the facts of arithmetic than to teach all the words in the English language, or all the topography of the globe, or all the details of human physiology.' The second is—'It is not desirable to eliminate any element of arithmetical training until you have something better to put in its place'."

The volume is replete with stimulating suggestions. Professor Thorndike has added a very concrete contribution to the better teaching of arithmetic as well as to teaching in general. It may well be read by the classroom teacher.

CYRUS D. MEAD

University of California

KUHLMANN, F. *Handbook of mental tests.* Baltimore: Warwick and York, 1922. 208 pp.

This handbook presents a new revision and extension of the Binet-Simon intelligence scale. The revised scale is given to the public after seven years of continuous work on the part of the writer, Director of the Minnesota Research Bureau in Mental Deficiency, and a number of assistants. The chief emphasis of the book is directed toward the presentation of the system of mental tests; but many important items of information in regard to the general principles of the year scale and the conduct of a mental examination are included, and a table of intelligence quotients is appended. Since, valuable as these latter features are, the claim of the book upon public attention will be based primarily upon the contribution which the new revision makes to the field of intelligence measurement, the review is concerned primarily with a comparison of the revision with the Binet-Simon scale and with other revisions.

This comparison touches upon the following points: (1) total number of tests included in the scale; (2) number which are reproductions or modifications of Binet tests or tests of other revisions; (3) number shifted from the age-groups assigned in other scales; (4) number of new tests; (5) range of the scale; (6) general procedure in the administration of the tests, nature of the instructions to the subject, and method of scoring; and (7) number of subjects used in the standardization. The scales compared are the Binet-Simon scale, 1911 edition, the Kuhlmann 1912 revision, and the Stanford revision.

The new Kuhlmann revision comprises 129 tests (including repetitions), as compared with the 54 of the Binet-Simon 1911 edition, the 75 of the Kuhlmann 1912 revision, and the 90 of the Stanford revision. Of these, 37 are found in similar or identical form in the Binet 1911 edition, 61 in the earlier Kuhlmann revision, and 45 in the Stanford revision. Thirty-five tests are common to all four scales. Only 5 of the 45 tests common to the Stanford revision are found in neither of the other two scales compared. Sixteen tests are shifted from the position which they occupied in the Binet revision, 9 from that occupied in the earlier Kuhlmann revision, and 14 from the position assigned by Terman. The number of new tests is 60 if the first appearance of a test as well as each repetition is counted; or 23 if repetitions are not counted—approximately the same as those added in the Kuhlmann and Terman revisions considered.

When the 23 new tests are analyzed, it is found that a strikingly large proportion involve some use of geometrical forms, considered in detail, they require making simple learned responses, recognizing objects, naming from memory, tracing, dealing with geometrical forms, counting, following directions, imitating folding, giving word opposites, spelling words backwards, crossing out letters in a pied text, working arithmetical problems, and making logical inferences. The range is from three months to fifteen years, as compared with a range from three years to adult in the Binet series, from three months to fifteen years in the earlier Kuhlmann revision, and from three years to superior adult in the Stanford.

In the general procedure in the administration of the tests, the instructions to the subject, and the scoring of the responses, the old tests appear to be patterned, almost without exception, directly after the earlier Kuhlmann revision, the text in instances being a verbatim reproduction. The resemblance to the text of the Stanford revision is also very striking, due to the common origin of the tests, although, as a rule, Terman's presentation is considerably more detailed. The new tests are very much like the older tests so far as their administration and the nature of the instructions are concerned; they differ markedly in the scoring, however, practically all those added for the years above two being scored in terms of the time required to perform the test and the number of errors made. The standardization of the scale is made as a result of tests with over 2000 normal children, in comparison with the 200 normal children used in the arrangement of tests for the Binet scale, the indefinite number employed in the earlier Kuhlmann revision (available data on the subject from other investigators having been included as well as that secured by Kuhlmann himself), and the approximately 1000 unselected children and 380 adults used by Terman.

In view of these comparisons, it is evident that the new revision is built primarily upon the Binet tests and the author's earlier revision, and in but small degree upon the Stanford. As compared with the first two scales, it is doubtless highly superior; as compared with the Stanford, it may or may not mark an advance.

Points which must arouse unfavorable criticism are the lack in objectivity of the infant tests and the inadequacy of their standardization, failure to include tests for the measurement of the intelligence of adults of superior mental capacity, the large quantity of test material required, and the elaborateness of the scoring formulae and the procedures for the computation of the mental age. Topics which will be subjects of controversy are the accuracy of the new positions assigned the tests shifted from the age-groups in which they were placed by other investigators, the validity of the new tests, the frequency with which the same test is used in different age-groups, the conception that the rate of mental growth decreases with increase in age, and the use of the results of examinations only of normal children in the establishment of norms. Admirable features of the scale are the simplification and improvement of the technique employed in the administration of some of the older tests, the elimination of some tests too largely dependent in their scoring upon the personal equation of the examiner, the introduction of others the scoring of which is almost wholly objective in character, the provision for the computation of a satisfactory mental age for those who pass more than two tests in the highest age-group, and the inclusion of a table of intelligence quotients.

Final evaluation will have to await wide use of the new scale by competent examiners and the determination of its validity in terms of its correlations with other criteria of intelligence. It will also have to await the determination of the comparative validity of such remarkably simple and convenient scales as the Herring revision of the Binet-Simon tests, and the various group tests for the measurement of intelligence.

LAURA M. CHASSELL

Ohio State University

DEWEY, EVELYN. *The Dalton laboratory plan.* New York: E. P. Dutton and Company. 1922. 173 pp.

Miss Dewey who has for many months been in close touch with Miss Helen Parkhurst, the originator of the Dalton Laboratory Plan, describes this experiment in school organization.

The fullest establishment of Miss Parkhurst's ideas is in her own elementary school, the Children's University School, in New York City. There the children, beginning with the fourth grade, spend their mornings from 8:30 to 12 in "laboratory" work. The hour from twelve to one is occupied with class conferences and auditorium exercises. The afternoon is given over to work in special subjects. In the three and one-half hours of each morning a child is free to work on the subject or subjects which he wishes—and in whatever order he chooses for as long a time as he thinks best.

Two factors in the organization which make possible this freedom are the "subject laboratories" and the monthly "contracts" of work in all subjects. Each of the five academic subjects in the curriculum is to be found in its own laboratory—a room in which all the teaching and studying in that subject is done. When a child goes into a laboratory he finds there the subject teacher and numbers of other children all engaged in work on the same subject, though they may be doing any phase of it from fourth-grade to eighth-grade complexity. It is a rule of the laboratory that a child may cooperate with any others of his own grade who are in the laboratory, help younger children, or ask for assistance from older pupils. The social experience within the child's own group and with other grade-groups, Miss Parkhurst considers two of the important aspects of her "laboratory" plan.

At the beginning of each month assignments in the subject for each grade are posted in the laboratories. The child copies the assignments on his record card, and this constitutes his contract for the month. The assignments in the subjects are arranged in blocks which are equivalent to one or two day's work, so that the child may enter on his record card how much he has accomplished in his morning's work in a given laboratory.

Children develop various types of procedure. Some children prefer to finish all the assignments in one subject and then go on to another. Other children do a little on each subject each day. How they may apportion their time is a matter left entirely to the children's discretion. Practice in controlling the day's activities is the unique contribution which the laboratory plan makes to children's educational experience.

Miss Dewey's account of the plan includes the detailed arrangements for the teacher's supervision of the child, the keeping of records, the making of assignments, the nature of assignments, and the like.

To quote Miss Dewey in her summary of the book—"Complete experiences, individual work, and scope for interests are essential in both the Dalton plan and the project method; when either is explained in theoretical terms.... The project method has freed itself from one side of this conception, the school of classified and isolated facts. The Dalton plan frees itself from the other half, the school of piecemeal assignments, bells, and herd-learning and recitations. One contributes a new subject matter to meet the needs of modern life; the other a way to give children working conditions that accord with the discoveries of modern psychology." HARRIET O'SHEA

Bryn Mawr College.

VOELKER, PAUL FREDERICK. *The function of ideals and attitudes in social education.* Columbia University, Contributions to Education No. 112. New York: Teachers College, Columbia University, 1921. 126 pp.

"You can not measure the human soul." A decade ago such was the slogan of the conservatives in their opposition to educational measurements, The slogan implied in particular that it is impossible to measure achievement or improvement in such school subjects as handwriting and arithmetic. But the events of the past ten years have shown that the conservatives were mistaken. Numerous tests are now available for use in practically all of the elementary-school subjects, and there are hundreds of practical school people who are ready to testify that the tests are worth while. Marked success has also been attained in the measurement of intelligence. Now Dr. Voelker comes with a book in which he undertakes to measure the affective side of mental life. The contribution is pioneer and even epoch-making in character.

The organization of the book is logical. The first chapter is entitled "The Perspective." The second chapter deals with the objectives of "Social Education." The third is devoted to the "Problem of Methods," the fourth to the "Problem of the Function of Ideals". The remainder of the treatise deals with the data obtained in an experimental manner.

In his second chapter the author defines his problem and marks out some of the chief objectives toward which social education should be directed. As such he specifically discusses the ideal of trust-worthiness, loyalty, social service, social sympathy, social conscience, social cooperation, social initiative, social justice, social control, tolerance, reverence, and faith. While some may prefer to state matters differently, all will agree that every effort to promote a clear conception of the objectives of social education is worth while. Definiteness of aim is nowhere more valuable than in the realm of morals.

Dr. Voelker's work would have been valuable even if he had stopped at this point. He goes further however. Having established his objectives he proceeds to a consideration of the problem of "Methods in Social Education." The essentials of such methods are stated in twelve assumptions which constitute the most complete summary of the sort which has yet appeared.

The author's most important contribution however, is in his last chapter in which he supports his theses by evidence, derived from actual experimentation. Having decided to make the ideal of trustworthiness the objective of the training, two series of tests of trustworthiness were devised which are as nearly alike in character as possible but quite different in content. Series 1 was given to four groups of boys and the scores were obtained. Then two of the groups received training in trustworthiness by methods "such as are commonly used by the leaders of the Boy Scout Troops;" the other two constituting control groups. After seven weeks of training the second series of tests were given. The scores were tabulated and compared with those of the first test. The two experimental groups had improved 13.5 per cent and 9.9 percent respectively while both control groups had failed to improve.

The elements of the experiment which are both unique and significant are these. The author has shown quite clearly that it is possible to provide a type of training which will strengthen and enlarge the ideals of the pupil. He also shows that improvement obtained from training, which is general in character can spread into various specific cases in which a general ideal is supposed to function. Tests have been very ingeniously devised which measure such an abstract and intangible quality as trustworthiness. A parallel series was provided so that improvement in the quality could be measured. Thus the way has been broken for the measurement of the affective phases of mental life.

Future investigations will bring many refinements of technique. Better methods of scoring, more adequate criteria and greater reliability and validity of the measures will doubtless be secured but this study will remain conspicuous as one of the first successful efforts to measure abstract personal qualities of the affective type by objective standards.

W. J. OSBURN

Department of Public Instruction, Madison, Wisconsin

NIEMEYER, N. *Children and childhood.* New York: Oxford University Press, 1921. 206 pp.

The purpose of the author is to give a cross-section of the child at various stages from birth to sixteen years. The reader is not to consider any characteristic peculiar to one period, but at a particular age, some given characteristic may stand out more than at another. The life of child is divided into the following periods: 1 to 5, 5 to 7, 6½ to 8½, 9 to 11 or 12, 12 to 16. The first is characterized by the education of muscles and nerves. The predominant change is the establishment of nervous connections in the cerebral cortex. Lack of nervous connections accounts for slowness of reaction on the part of the very young. For developing these, the child needs freedom of movement, nourishment, large playthings, and small things to use. Proper brain connections should be established by the formation of habits. The diagrams used to illustrate this leave much to be desired in respect to accuracy.

When the child commences to classify objects, however crudely, he has passed the first stage of babyhood. Imagery and imagination now stand out prominently in his play. In these he differs from his younger or older brothers and sisters, for the younger ones enjoy play for sensory repetition, while the older ones desire more reality. Emotions of self and sex are also becoming evident. The writer agrees heartily with Miss Niemeyer in the fact that children should be taught simply but accurately facts about sex. Evasions lead to false suppositions. This is also the period of instilling proper habits of discipline and thinking. Discipline should not be repressive.

The period of 6½ to 8½ years is one of transition. The child differentiates between the real and fantastic. Dullness and backwardness, if present, are more evident than previously. Between the age of 9 and 12, the child shows more stability, physically and mentally, than in babyhood. During this period especial attention should be given to the training of the memory. Miss Niemeyer also emphasizes the necessity of transfer of habits at this age by producing classroom situations more like those of the world, for only when the two situations are alike in one essential point is there transfer. She hastens to explain in a footnote that this really signifies no transfer. The child's thought and interest are self-centered. Attention to these, however, does not have to be created, but has to be trained.

The last period, or that of adolescence is marked by the arousal of sex- and self-emotions. Self-emotion leads to reflective thought and self-willed actions. During this age, the adolescent should be taught to decide his own problems of discipline. All of these factors properly guided should lead to the establishment of a strong character.

Miss Niemeyer's book is well worth the time of perusal of the mother interested in the psychology of her child or the student beginning the study of child life. It would serve as a good introduction to the possibilities of the subject. But for the more mature student, the book is scarcely thorough enough. One of the best features of the book is the original excerpts appended to illustrate various characteristics of the different ages of childhood. After reading this little volume one should understand children better in the light of their activities and conversation; and that, after all, was the main purpose of the author.

A. Sophie Rogers

Ohio State University

Bamberger, Florence Eilau. *The effect of the physical make-up of a book upon children's selection.* Baltimore: The John Hopkins Press, 1922. 161 pp.

The monograph presents the results of an experimental study of the factors which influence children's selection of books for reading. The desirability of such empirical approach is brought out for it develops that adults' judgments regarding the books which children will choose disagree markedly with the actual findings.

The material consisted of five familiar children's stories each in five editions. These are described in detail and apparently differed widely in physical properties—external and internal. Each story in its five editions constituted a separate experiment. The story was read by the experimenter up to a certain point of suspense and the children were asked if they would like to finish reading it themselves. The five editions containing that story were then placed upright on the chalk rail of the blackboard and a number written above each. The space error was obviated by altering the arrangement of the books with the different groups of subjects. The children came forward one at a time, inspected the books without touching them and recorded on a slip of paper tho number of the book in which they wished to finish the story. Each child was later given the five editions, was allowed to examine them freely and was then asked to indicate the one preferred. If this choice differed from his first (''inspection'') choice he was asked the reason and his comments were noted. The experiments were performed with groups of about ten children at a time. The teachers acted as experimenters and recorded the original results. The subjects were pupils in the first three grades of five widely differing schools in New York and Maryland. Each series involved results from 76 to 260 subjects.

The results of the experiments are presented in great detail with tables and charts. The size of the book seems to be a significant factor in selection. This result is further corroborated by additional experiments in which all the books were covered with tan paper. This same experiment shows that the appearance of the cover also exerts an influence. Both the color of the cover and the nature of the title are involved. It is interesting to note, with the older children, the negative effect of the word ''reader'' in the title. The outstanding factor influencing choice, after thorough examination, is illustration. In the books that were chosen more than one-fifth of the time (chance expectation) the average amount of space devoted to illustrations is 45 percent while in the books that were chosen less than one-fifth of the time the average amount of space devoted to illustrations is 27 percent. There seems some evidence that illustrations portraying humor or action and having story telling qualities are preferred. Few lines to a page and wide margins seem desirable. The fact that in 45 percent of the cases the books were retained by the children after an opportunity to examine them, points to the benefits of open shelves in libraries. Differences in sex and school grade throughout the experiment are slight.

The monograph concludes with a score card for evaluating the various factors in the physical make-up of a book for primary children. Each of a large number of factors is weighted according to its actual frequency of appearance in some 2000 choices. Then fourteen of the books used are evaluated according to this scale. Their values range from 76 to 38 and correlate perfectly with their actual frequency of choice.

The method of the experiment seems sound and if strictly adhered to should yield results void of suggestion and space errors. The number of subjects seems sufficient to afford valid conclusions. The weakest part of the experiment in the reviewer's opinion is the material, involving as it does so many simultaneous variables.

The reviewer intends by no means to belittle the monograph. It is a distinct experimental contribution in a field that has not been hitherto approached in this manner. The method is novel, the subjects adequate and the presentation of results detailed and lucid. It points the way to further study of certain variables which did not come out clearly with the present technique but which are perhaps of considerable significance.

HAROLD E. BURTT

Ohio State University

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors. Address all correspondence concerning this department to Doctor E. J. Ashbaugh, Ohio State University, Columbus, Ohio.

Doctor John L. Stenquist, assistant director of the Bureau of Reference, Research, and Statistics at New York City, has been appointed Director of Educational Research for the public schools of Baltimore, Maryland.

Doctor Leo J. Brueckner, assistant dean of Detroit Teachers College has accepted an assistant professorship in the College of Education, University of Minnesota, beginning with the current semester.

Doctor Marion R. Trabue, head of the Bureau of Educational Service, Teachers College, Columbia University, has resigned that position to become Professor of School Administration in the University of North Carolina.

At Stanford University, Dr. Lewis M. Terman has been appointed head of the department of psychology to succeed Professor Frank Angell who retires at the end of the academic year 1921-1922. Other additions to the department include Dr. W. R. Miles, professor, and Dr. Calvin P. Stone, assistant professor.

Doctor William A. McCall, assistant professor of school administration, Teachers College, Columbia University, is spending the current year in China on invitation of the National Association for the Promotion of Education to assist in developing scientific education through the creating and adapting of tests, formulating experimental and statistical methodology, and advising about researches.

The Educational Research Department of the Ohio State Teachers' Association met at the Hotel Breakers, Cedar Point, Wednesday evening, June 28, 1922. Some months ago the members voted by mail to dispense with formal programs at the summer meetings but to hold a dinner, followed by an after-dinner address. The meeting was in charge of the vice-president, Doctor A. R. Mead of Ohio Wesleyan University. Doctor E. J. Ashbaugh of Ohio State University addressed the meeting upon the theme ''Some Essentials in Supervision.'' The address was well presented and well received. The organization instructed the secretary, Superintendent F. J. Prout of Sandusky, to mimeograph the address and mail copies to all public school superintendents in the state as well as to the members of the department. This was done within two days of the date of the meeting.

Our readers will be interested to know that the Charters Diagnostic Language and Grammar Tests will be made more useful this year than they have ever been before. We have all recognized the fundamental soundness of Doctor Charters' language and grammar studies—studies based on the errors which children make. Unfortunately the particular form of the tests has not hitherto exhibited the best technique, the result being that the tests have fallen short of the objectivity which modern standards require. The new forms of the tests will be ready for distribution in September. With the cooperation of a number of superintendents the new edition is being standardized so that tentative norms will also be available this fall. The tests will continue to be published by the Public School Publishing Company.

We recently received a brief letter which, with the omission of the name of the city, read as follows: "I am very happy to tell you that we have no supervising principal in ————. In my opinion that would be the greatest non-essential that could be added to the system." We naturally wonder concerning the value of this opinion. Has the writer scientifically evaluated the work of supervising principals? If so, and the results have shown them to be non-essential, then this is not an opinion but a statement of fact. On the other hand, if such an evaluation has not been made, from what source has the writer derived the right to express an opinion upon the subject?

The London Teacher for July 7, 1922, calls attention both by editorial comment and news item to an action taken by the education committee of the London County Council which will be of great interest to educators everywhere. The report contains a very significant statement, namely, "we have considered this question very carefully and are of the opinion that the children would not suffer if the classes mainly composed of children under five were taught by the same type of women as was employed for this purpose during the war." The type of teacher referred to was the emergency teacher who had received a three-months intensive course of preparation for teaching very young children. The argument for this action is an economic one, the difference in salary between a certified teacher and an uncertified one being about one hundred five pounds per year, or something less than five hundred dollars. An experimental arrangement with a group of one hundred young teachers is to be attempted. The question is raised editorially and we may well repeat it in this country as to whether it is economical for unprepared teachers to teach children of any age.

The following expression of attitude adopted by the New York Federation of Labor in its fifty-eighth annual convention this past June is worthy of the most careful consideration of school people everywhere.

"School statistics show that the great mass of children leave school between 14 and 16 years of age, and that the eighth grade or lower is the scholarship attained by this group. Many parents and children regard graduation from the VIII-B grade as the end of education; and this impression is heightened by the elaborate public ceremonies attending graduation exercises when some overzealous trustee or other prominent citizen attempts to advise the graduates on the brilliant careers open to them. There should be no break in the continuity of advancement to the higher grades, and *graduation exercises should only be held on completion of the fourth year in high school.*''

The chairman of the committee on education, in forwarding this statement to the school people of the state, recommended the abolition of elaborate graduation exercises and the discouragement of the tendency toward elaborate apparel for graduates.

Rochester, New York, has been one of the leading cities of the United States in the establishment of junior high school. Four junior high schools are now in operation; and by 1925 it is expected that the entire city will be on a 6-3-3 basis. The development of the work has been gradual, and the Municipal University has devoted part of its energy to the preparation of teachers for this type of work. The salaries in junior high schools are the same as in senior high schools, and the teachers must have had two years of normal-school training, one year of college work, and three years of experience.

The first junior high school in the city was organized in 1915 in the Washington school district which has more than 90 percent foreign population. The average carry-over to the ninth year for five semesters previous to the opening of the junior high was 58.7 percent. The carry-over for the succeeding ten semesters averaged 85.8 percent. In comparison with this, the carry-over in the Madison school district, which is a representative American section, was 73.6 percent and 78.2 percent for the corresponding periods. Such information presents real evidence of the effect of the junior high.

Assistant Superintendent J. P. O'Hern, who has furnished us with these data, further states: ''The junior high school is a great factor in Americanization work. In the Washington section the Jews would remain in No. 9 school, the Italians in Nos. 18 and 27, the Poles in 22, the Germans in 26, etc., until they graduated from the eighth grade. Now all the foreigners enter the junior high school before the law allows them to leave school. They touch elbows, become friends, and consequently get rid of their natural prejudices and hatreds. The junior high is democratic in offering facilities according to needs and abilities and in bringing pupils together at a time when friendships are formed and prejudices removed.''

The following communication from Dr. Frank N. Freeman of the University of Chicago, concerning some of the work in the educational laboratory there, will be of interest to those who are interested in the development and use of refined technique in the evaluation of educational procedure.

''The investigations of the mental and physical development of children have been conducted for the most part by making measurements or tests of different children of different ages. The averages of the successive ages have then been used to construct a curve of development which has been regarded as the norm. We have no assurance, however, that such a curve represents at all faithfully the course of development of individual children.

In fact the few studies in which repeated measurements have been made in succeeding years upon the same children have indicated that individuals deviate to an important degree from such averages. In the second place it has not been usual to make a variety of both physical and mental tests of the same children in order that the relationship between the different aspects of the child's development may be correlated. There have been, it is true, exceptions to both of these generalizations. It is desirable, however, that a much more comprehensive, systematic, and long-time study of the development of a group of children be made than has been made up to the present time. Such a study is being initiated through the cooperation of the educational laboratory and the laboratory schools in the University of Chicago. The plan contemplates keeping record of the physical measurements including X-ray photographs of the bones, making physical and psychophysical tests, and making a record of educational achievement and of changes in emotional attitudes on the same children for a period of ten years, or as long as the children remain in the schools. It will be possible to secure from time to time comparisons between the results of these various tests, but the fundamental object of the study will be to trace the various forms of development as they appear in the same individual children and to study their relationship to one another throughout the period of the investigation.''

Inferences from Small Samples and Differences

In a recent article[1] in this Journal, Mr. J. W. Heckert has presented an excellent plan for determining the value of supervised study in English composition. The improvement of supervised over unsupervised groups is shown by the gains over preliminary test scores on the Thorndike-Hillegas Scale. Two other types of comparison are made by separating the groups into halves according to intelligence (National Intelligence Test), and according to initial test score. From all of these plans the writer finds differences in favor of the supervised group, and concludes that, ''Supervised study in English composition is eminently worth while when teachers are able to direct the children's efforts intelligently.''

While the plan of this study is admirable, the conclusions from the results obtained are open to question because of the small number of cases involved (8, 16, 17), and the slight differences in gain. According to Table I, the average gain for the supervised group is 2.6 percent. The question arises as to whether the difference of 4.4 percent is really significant or not. On this point the validity of the whole conclusion rests. The obvious answer is to determine the probable error of the difference by the formula

$$P.E._{a-b} = \sqrt{P.E._a^2 + P.E._b^2}$$

---

[1] Heckert, J. W., ''The effects of supervised study in English composition,'' *Journal of Educational Research*, 5: 368-80, May, 1922.

The quantities $P.E._a$ and $P.E._b$ are obtained for the respective mean gains from the formula

$$P.E.\ \text{mean} = .67449\ \frac{S.D.\ \text{distribution}}{\sqrt{N}}$$

giving

$$7.0 \pm 1.5 \text{ and } 2.6 \pm 1.4$$

The probable error of the difference between the two means is then

$$\text{P.E. difference} = \sqrt{(1.5)^2 + (1.4)^2 = 2.1}$$

Therefore difference $= 4.4 \pm 2.1$

By making use of Pearson's Tables of the Probability Integral we find that the odds are roughly 5 to 1 against these two means being identical, or merely random samples from the same population, or that the difference is really significant. As a matter of fact the probable errors for such small samples are even larger than those given by the above formulae, and the chances for significance of the differences even smaller. The usual rule is that the observed difference must be three or four times its probable error in order to be significant. This places the odds at roughly 100 to 1 against the difference being due to sampling.

Differences and their probable errors have been calculated for the remainder of the writer's results and are given in the table below. Inspection of this table shows that none of the differences are truly significant, the greatest odds in favor being roughly 12 to 1. These are hardly safe betting odds on which to base scientific conclusions. It is interesting to note, however, that the most reliable results revealed by the table are that the brighter pupils profit more by supervised than by unsupervised study, and that the initially poorer pupils on the test showed the greater improvement. This suggests that the bright lazy boy will profit under supervision, while the dull hard worker will not, a result often checked by experience.

MEAN GAINS AND DIFFERENCES FOR SUPERVISED AND UNSUPERVISED GROUPS

GROUP	SUPERVISED		UNSUPERVISED		DIFFERENCE	ODDS FOR SIGNIFICANCE
	Mean	S. D.	Mean	S. D.		
All.........	7.0±1.5	8.8	2.6±1.4	8.5	4.4±2.1	5:1
Brighter half	9.8±2.4	10.0	1.7±1.9	8.1	8.1±3.1	12:1
Duller half..	4.3±1.5	6.4	2.7±2.2	9.3	1.6±2.7	....
Better half...	3.7±2.2	9.3	0.2±2.0	8.4	3.5±3.0	....
Poorer half..	10.4±1.6	6.8	4.1±1.8	8.1	5.8±2.4	9:1

It is to be hoped that the above experiment will be repeated with larger numbers, or adapted so as to reveal greater differences if they exist. If it seems inadvisable to use larger groups or different methods of supervision, a longer training period might yield significant results.

KARL J. HOLZINGER

University of Chicago

Intelligence Tests and College Success

During the first full week of the present academic year four different group-intelligence tests were administered to the one hundred new students entering Hood College for the first time. One test was given on each of four successive mornings. The four tests used were: Terman, Otis, Thurstone, and Rogers.

A study of the scores seems to justify two conclusions:

1. *That the four tests measure the same thing, whatever that may be.* In support of this first conclusion the following correlations are given (the method used is that of the percent of unlike-signed pairs):

	Coefficient of correlation	*P. E.*
Terman and Otis	.90	.01
Terman and Thurstone	.86	.02
Terman and Rogers	.61	.04
Otis and Thurstone	.90	.01
Otis and Rogers	.66	.04
Thurstone and Rogers	.73	.03

The following correlations were obtained between each test and the average of the four tests:

	Coefficient of correlation	*P. E.*
Terman	.93	.01
Otis	.97	.005
Thurstone	.96	.005
Rogers	.83	.02

2. *That the thing measured by these tests in this case was not the thing which determined success in college.* In other words, the scores derived from these tests would not have furnished a satisfactory basis for predictions as to the probable success of individual students in the college work of the first semester. For when we correlate the intelligence scores with the average college grades we get the following results:

	Correlation with college success	*P. E.*
Terman	.48	.05
Otis	.06	.07
Thurstone	.125	.065
Rogers	.34	.06
Average of the four tests	.45	.05

We undertook this study with a prejudice in favor of the tests and with the belief that the average scores secured from these four tests would furnish a much better basis for predicting college success than appears to have been the case. The coefficients in the case of the Terman and Rogers tests are good enough to be tantalizing. They also arouse the hope and justify the belief that it should soon be possible to improve group-intelligence testing to the point where we shall get correlations with college success that will run well above 0.60.

In this case the tests have done *something* with beautiful consistency, but that thing is not just exactly the thing that we desired to have done.

Hood College, Frederick, Maryland

Elmer Hoke

A Revision of the Pressey Primer and Cross-Out Scales

Four years ago the writers developed two "brief group scales for measuring general ability for use in school surveys," the *Primer* Scale and the *Cross-out* Scale. The examinations were the result of a special effort at construction of scales which should be easy to give and score, inexpensive, and readily handled. They involved many radical departures from current practices in test building. In consequence, they were offered to the educational public with some misgivings as to the reception they might be accorded. Their wide-spread use, since that time, would indicate that they filled a real need. However, numerous short-comings have also appeared. It has therefore seemed advisable to attempt a revision of the series—a revision which this communication aims briefly to summarize.

The revision of the Primer scale has aimed primarily at better adaptation to work with first-grade children. One test of the first scale (the form or block test) has been entirely eliminated, as too difficult to explain to young children. The revised blank shows on the first page a directions test. This serves not merely as a test; it also accustoms the children to the procedure of the examination—habituates them to moving across the page from left to right (something many first-grade children have not yet learned), and so on. After taking the first test the children simply turn over the test folder instead of opening it; the second test is a revised pictorial absurdities test. The blank is next opened to an improved pictorial selection test showing four objects to each picture instead of three, in order to cut down the possibility of chance success. A revised and improved "dot pattern" test comes last. That the scale has in fact been made much more applicable to the first grade may be appreciated from the fact that, in a large school recently surveyed by the writers, no I-B child scored below 8 points. The scale is even more easily given and scored than the original Primer scale. More extensive norms than are available for the first form are now available for this revision.

The first Primer scale was found not entirely satisfactory for the third grade; third grade retards scored high, as a result merely of maturity. So the new Primer scale has been designed for use in the first two grades only. And an "Intermediate Classification Test" has been devised, to cover grades II-VI inclusive. The scale can perhaps hardly be called a "revision" of the Cross-out scale. But it is the direct result of the experience gained with that scale. It is similarly compact, and even more easy to use. On the first page of the small four-page folder appear the directions for all four tests of the scale, with examples from each. On the succeeding three pages are the 96 questions of the test proper, the four tests being arranged in cycle-omnibus form. The directions for the entire scale are thus given once for all at the beginning of the examination period; after this the examiner has nothing to do but collect the blanks when the time limit is reached.

The "Senior Classification Test" for Grades VII-XII is similar in organization to the intermediate scale; but the questions are, of course, of a more advanced and difficult nature. By thus covering with two scales the range of grades previously covered by the Cross-out scale it has been possible to secure a much better differentiation; the grade over-lapping, for instance, is much less.

It is believed that this series of scales will prove of distinct value as brief instruments, for measuring general ability, for use in school surveys, and in other projects where an inexpensive, easily used, compact examination is desired. It may be added in closing that complementary examinations for each scale are contemplated, after the fashion of the Scale A and Scale B of the National Tests; when more reliable ratings on individual cases are desired it will then be possible to give the second form on a succeeding day. The total rating will then be much freer from the influence of chance circumstances than the findings from a single long examination.

S. L. Pressey
Ohio State University L. C. Pressey

Superintendent W. W. Lewton of Cicero, Illinois, informs us that his school system has been working on the "project" for the past three years. They have experimented, weighed results, and studied all the literature they could find on the subject. He feels that they are now seeing light both as to procedure and content.

The following words from a letter are self-explanatory: "We have found some big hurdles that must be cleared. First, the supervisors must be extremely well informed and they must be expert salesmen. Second, the proposition has to be sold to the teacher because the procedure is such a radical about-face in schoolroom activity that the 'informational' type of teacher, as well as the one deficient in leadership and initiative, will never be able to obtain results. Third, printed material, outlines, and suggestions are not extant. Fourth, our effort is pioneer work. We find it so because the 'problem project' is such a catchy phrase that many people have looked upon it as a panacea for all school ills."

We hope that Superintendent Lewton will organize the material which he has been gathering in such a way that it may be presented to the educational public.

Some pertinent sentences in the Bulletin on the Measurement of Ability in Spelling by the Bureau of Reference, Research, and Statistics of New York City are worth passing on for the thoughtful consideration of our readers. Read them carefully, and consider the extent to which you are applying the truths stated therein in your own school.

Improve your instruction by means of standard tests and measurements. The setting up of educational standards has resulted in the elimination of useless subject-matter. Economy demands that no more time be given to a subject than is necessary for the pupils to attain a satisfactory standard of achievement. Such standards are provided by educational measurements. Simply to know a pupil is below standard in ability is of little value to the teacher, because spelling ability is specific and not general. It is important that the teacher think of tests and scales as instruments which will enable her to make her instructional efforts more effective. Drill to be efficient, must be made individual in character.

Educational Research Association

(Frank N. Freeman, *Secretary and Editor*)

An Analytical Study of the Results of the Woody Arithmetic Test in West Allis, Wisconsin

Communicated by Mr. T. L. Torgerson,
Department of Educational Measurements

The Woody Arithmetic Scales were given to 1282 pupils in grades III to VII. The tabulation of the scores indicated that the pupils were up to or above the norm. The feature of the test which is of most interest to us, however, is the use which was made of the results to determine the frequency of error and the causes of these errors. The report (a part of which is here summarized) contains a considerable list of errors for each operation and the frequency with which these errors occurred. Space will not permit the printing of the full list. Table I, however, gives a summary of data for the five most frequent errors in each operation:

TABLE I. THE FIVE MOST PREVALENT TYPES OF ERROR IN EACH OPERATION

Error Number	Problem Number	Incorrect answers obtained	Frequency in each grade (percents)				
			III	IV	V	VI	VII
		Addition					
1	35	9 ft. 20 in.			15.0	33.9	31.5
2	36	20 yr. 29 mo.			9.1	33.6	22.9
3	7	3	11.9	10.7	9.9	6.2	2.2
4	27	3/14		1.1	5.2	15.5	5.7
5	24	183762		2.1	5.0	4.2	10.5
		Subtraction					
1	25	15 5/8			30.5	26.1	11.0
2	27	2 yd. 8 ft. 6 in.		4.2	26.5	24.5	13.0
3	35	2 2/8		6.4	21.0	21.2	18.6
4	14	35	20.3	14.0	6.6	2.4	
5	14	1 yd. 9 ft. 3 in.	21.7	13.2	4.6	2.8	
6	26	30		3.2	28.6	33.6	13.5
		Multiplication					
1	34	1			10.1	26.8	11.7
2	35	24450 ¾			10.1	20.0	1.3
3	21	48 ½		4.2	16.9	9.6	2.0
4	19	384	4.2	3.9	14.9	2.4	4.8
5	22	40 ¾		4.2	13.2	12.0	1.3

Division

1	35	71 21/531			.5	20.4	26.0
2	31	¾			1.7	26.1	16.9
3	26	23		.7	8.9	26.1	.9
4	22	675		5.6	24.5	19.6	5.8
5	33	72 12/52			2.6	22.0	13.5

TABLE II. CAUSES OF ERRORS

Error Number	Cause
	Addition
1.	Failure to reduce to higher denomination in denominate numbers.
2.	The same as 1.
3.	Multiplication for addition.
4.	In fractions, adding the numerators for a new numerator and the denominators for a new denominator.
5.	Failure to use the decimal point.
	Subtraction
1.	Bringing down the fraction from a mixed number subtrahend.
2.	Ignorance of table of linear measure.
3.	Failure to reduce answer to lowest terms.
4.	Failure to pay back.
5.	Failure to borrow.
6.	Ignorance of table of linear measure.
	Multiplication
1.	Addition for multiplication.
2.	Failure to multiply by the fractional part of a mixed number.
3.	In mixed numbers, multiplying integers and adding fractions.
4.	Decimal point placed wrong.
5.	Failure to multiply by the fractional part of a mixed number.
	Division
1.	$a \times o = a$.
2.	Failure to place decimal point.
3.	Errors whose cause could not be discovered.
4.	The same as 3.
5.	$a \times o = a$.

From the data which are given elsewhere in the report it is possible to determine the causes of these most frequent errors. They are given in Table II. A number of very interesting facts appear from this table. The first is that, even though the number of errors here represented is small, we find repetitions of the same cause. The same cause appears more than once in a single operation and is also responsible for the mistakes in different operations. For example, the repetition of the failure to place the decimal point properly is found in several groups. Another interesting fact is that the errors in general appear to be rather trivial. They do not seem to depend in some cases so much upon the failure to understand as upon mere carelessness. A further elaboration of this type of study would prove very valuable as a means of correcting or anticipating the errors which pupils make in their arithmetic.

THE RESULT OF A STATE-WIDE SURVEY OF SILENT READING IN IOWA

Communicated by H. A. Greene,
Extension Division, State University of Iowa

In the survey which is here reported, the Monroe Silent Reading Test was given to a large number of schools and systems in Iowa, and the results were tabulated by the Extension Division. In all, results were obtained from 232 school systems, some situated in rural districts, others in cities. The distribution of the contributing school systems according to size is given in Table I.

TABLE I. CONTRIBUTING SCHOOL SYSTEMS[a]

GROUP	POPULATION	NUMBER OF CITIES
I.......	Less than 1,000	55
II......	1,000 to 1,999	45
III.....	2,000 to 2,999	12
IV......	3,000 to 4,999	11
V.......	5,000 to 9,999	12
VI......	10,000 and above	12

[a] In addition to the city systems listed in this table, 85 consolidated schools also contributed.

Approximately 110,000 copies of the Monroe Silent Reading tests, Form 1, were shipped to the cities cooperating. From the 147 cities and 85 consolidated school finally reporting, records for 75,276 pupils in grades III to XII inclusive were received. Table II shows the distribution of these pupils by grades and by size of city.

TABLE II. NUMBER OF SCHOOLS AND PUPILS

GRADE	CONSOLIDATED	GROUP I	GROUP II	GROUP III	GROUP IV	GROUP V	GROUP VI	TOTAL
III......	1,700	696	950	506	456	1,100	4,291	9,699
IV......	1,571	687	1,076	449	550	1,087	4,134	9,554
V.......	1,549	727	1,142	421	645	1,085	4,020	9,589
VI......	1,566	752	1,081	440	597	1,148	4,047	9,631
VII.....	1,521	765	1,173	447	586	1,014	3,607	9,113
VIII.....	1,532	721	1,252	420	554	875	3,300	8,654
IX......	1,005	706	853	233	286	876	2,726	6,685
X.......	799	536	762	112	225	666	1,798	4,908
XI......	627	429	669	125	249	622	1,425	4,149
XII......	506	321	575	109	179	531	1,073	3,294
Total...	12,376	6,340	9,533	3,272	4,327	9,004	30,424	75,276

The comprehension medians for each grade and group were computed from a composite distribution for each group of cities. That is, the scores made by the 1700 third-grade children from the 85 consolidated schools involved were distributed in a composite frequency table. From this composite table the median was calculated.

In the case of the rate medians lack of time prevented the making of similar composite distributions. Accordingly, in this case a weighted mean of the rate medians for each city group was found. This was checked sufficiently to show that the mean was fairly reliable and did not vary noticeably from a median calculated from a composite distribution. Evidence recently submitted with regard to the significance of Monroe rate scores leads to the conclusion that these rate scores are sufficiently accurate for ordinary purposes.

A comparison of the relative level of achievement of each grade of the different city groups with Monroe mid-year norms is shown in Table III. This table shows the percent that each grade median is of the mid-year norm. It is read as follows: the median score for third-grade pupils in the 85 consolidated schools is 67.5 percent of the Monroe norm for the middle of the year; the fourth-grade median for consolidated schools is 80 percent of the mid-year standard, etc.

TABLE III. PERCENT THAT GRADE MEDIANS ARE OF MONROE MID-YEAR NORMS

GRADE	CONSOLIDATED SCHOOLS	GROUP I	GROUP II	GROUP III	GROUP IV	GROUP V	GROUP VI	TOTAL	MONROE MID-YEAR NORMS
					Comprehension				
III	67.5	67.5	65.0	106.0	92.5	83.8	128.0	93.0	6.8
IV	80.0	78.5	93.5	101.0	102.0	85.8	112.0	97.5	12.7
V	75.3	82.5	83.5	95.5	87.5	92.0	104.0	94.0	17.8
VI	92.5	92.9	114.0	121.0	109.0	103.0	113.0	107.0	18.5
VII	88.0	88.0	102.0	109.0	106.0	98.0	107.0	101.0	22.8
VIII	86.0	84.5	86.5	104.0	112.0	99.8	103.0	96.5	26.0
IX	97.8	98.0	104.0	109.0	127.0	110.0	123.0	113.0	23.0
X	104.0	103.0	102.0	108.0	121.0	113.0	121.0	113.0	25.4
XI	104.0	104.0	112.0	104.5	123.0	124.0	123.0	116.0	27.2
XII	105.0	105.0	112.0	97.0	132.0	114.0	120.0	116.0	30.0
					Rate				
III	79.5	69.5	85.5	100.0	96.5	92.5	115.0	91.1	52.0
IV	97.0	87.5	91.5	101.0	100.0	85.5	111.0	96.5	70.0
V	90.5	92.0	93.0	108.0	110.0	101.0	107.0	101.0	87.0
VI	96.0	92.0	95.0	106.0	109.0	103.0	108.0	101.0	90.0
VII	92.7	92.1	94.1	109.0	110.0	98.6	105.0	100.0	106.0
VIII	88.3	92.0	103.0	104.0	115.0	99.0	109.0	101.0	83.0
IX	95.0	95.0	103.0	100.0	108.0	98.0	114.0	102.0	85.0
X	98.2	94.3	99.0	106.0	119.0	103.0	111.0	104.0	85.0
XI	98.5	93.5	100.0	104.0	111.0	105.0	106.0	103.0	90.0
XII	95.8	94.0	98.0	98.5	126.0	98.0	104.0	101.0	96.0

The average scores or median scores made by the various groups of pupils may be obtained by multiplying the Monroe norms in the last column by the percents given in the main body of the table.

From Table II it is clear that the Iowa results fall below the Monroe mid-year comprehension standards in grades III, IV, V, and VIII. This situation can scarcely be considered serious except in the case of the lower grades in certain of the small grade schools. It is significant that in practically all cases the composite results for the high-school grades lie well above the mid-year standards established by Monroe for these grades. It is also noticeable from these data that the Iowa results move forward uniformly from grade V to grade VI instead of showing the typical break in the curve as revealed by Monroe's results. As a matter of fact, from the Iowa results it appears that Monroe's norms for the sixth and ninth grades are considerably lower in proportion than those for the fifth and eighth grades respectively. It also appears that so far as rate of reading is measured by the Monroe test the schools tested in this survey are approximately up to standard achievement for the middle of the year.

The foregoing table offers some indication as to the relative level of silent-reading ability in the different classes of cities. It will be noticed that for cities of Group III the composite median for grades V and XII are the only ones which fall below the Monroe norms. In the case of Group VI no composite median falls below the norm. A consideration of these results seem to justify the conclusion that due to certain factors not identified in this report a higher quality of silent-reading comprehension has been developed in the larger school system.

The foregoing data indicate the level of achievement of classes in various types of city schools at the mid-year. In addition to these data, tables not reproduced here, have been constructed to show the large variability among the scores in individual grades. It is planned to give the test again at the end of the second semester. This will indicate the growth which has then taken place during a half year.

REQUEST FOR SUGGESTION ON ADMINISTRATIVE PROBLEMS

Assistant Superintendent Joseph P. O'Hern of Rochester, New York, wishes to secure information or suggestions on the following topics:

1. Promotion.—Based on class work, examination, or both; trial promotions and, if allowed, for what length of period; promotion based on course of study arranged for maximum, average, and minimum essentials; responsibility for non-promotion placed on receiving or sending teacher.

2. Per capita cost.—What items are figured into the cost, and what divisors are used.

3. The average number of pupils per teacher and for various types of schools.

4. Salary schedules for different types of service.

JOURNAL *of* EDUCATIONAL RESEARCH

Volume VI OCTOBER, 1922 Number 3

THE CLASSIFICATION BY TESTS OF INTELLIGENCE OF TEN THOUSAND FIRST-GRADE PUPILS

CHARLES S. BERRY
Director, Psychological Clinic, Detroit Public Schools,
Professor of Educational Psychology, University of Michigan

The plan of organization for the public schools of Detroit, adopted by the Board of Education in September, 1919, provides for the education of the pupils in the kindergarten and first six grades in large elementary schools with gymnasiums and auditoriums planned to meet the requirements of the platoon organization, and for the education of pupils in grades VII, VIII, and IX in intermediate schools large enough to accommodate from twelve to eighteen hundred pupils. "While it is important to provide for individual differences in the training of all children, it is doubly so during the adolescent period when these differences are more marked than they have been in the earlier years. A large intermediate school can provide for these differences by a varied curriculum, and by giving pupils some freedom to select types of work which they prefer."[1]

If the intermediate school is to fulfill its purpose in making provision for the adolescent, it is evident that some marked changes must be made in the classification and promotion of pupils in the first six grades, otherwise a large percent of these pupils will reach the age when they can leave school before they have completed or even entered the intermediate school. For the age-grade census of 1921-1922 shows that 17 percent of the pupils in the Detroit public schools fifteen years of age are either in the sixth grade or below, or in special classes; and that 22 percent of the pupils fourteen years of age are in the sixth grade or below, or in special classes. This means that a large percent of these two age-groups will never enter the intermediate school at all, since many of them will become sixteen years of age before completing the sixth grade. Yet these pupils are the ones that stand most in need of the kind of training and instruction offered by the intermediate school.

---

[1] Spain, Charles L. and others, *The intermediate school in Detroit.* (The Detroit Educational Bulletin. Research Bulletin No. 6.) Detroit: Board of Education, 1921, p. 8.

Under existing conditions we make it difficult for pupils of superior intelligence to develop the qualities of leadership by compelling them to associate with older pupils who resent being led by those who are younger; and we make it difficult for older pupils to select their leaders wisely by putting in with them those who are younger but brighter. It is true that in adult life age is a minor factor in the selection of leaders, but in childhood and youth age has much significance; for generally speaking, the older child is taller, heavier, stronger and of superior intelligence.

Education in a democracy should make adequate provision for the development into leaders of those who have the greatest possibilities of leadership, and at the same time, give training in the wise choice of leaders to those who are to follow. By the intermingling of pupils of about the same age on the playground, in the gymnasium, in the auditorium, in social and class activities, age ceases to be a factor in the development of leadership, or in the choice of leaders.

In order to correct, to some extent at least, the evils existing under the present plan of classification and promotion of pupils in the elementary grades, Superintendent Cody approved a plan during the winter of 1919-1920 for the classification by tests of intelligence of all pupils entering the first grade.

Because of the large number of pupils to be examined and the necessity for speed, if the results were to be available soon after the beginning of the semester when they would be of the greatest value, and because of the expense connected with the giving of an individual intelligence examination, it was decided to devise a group intelligence test for this purpose. The Detroit

First Grade Intelligence Test,[2] devised by the staff of the Psychological Clinic, to classify tentatively for purposes of instruction pupils entering the first grade, requires no knowledge of written English, and correlates .69 with the Stanford-Binet based on one hundred unselected cases.

The actual examining was done by the teachers of classes for subnormal pupils, assisted by the staff of the Psychological Clinic. These special teachers, who had previously been trained to give the Stanford-Binet, were specifically instructed in the method of giving the Detroit First Grade Intelligence Test.

During the first four days of school in September, 1920, more than ten thousand pupils in the B-1 were given this intelligence examination. The results of these examinations were sent to the principals of the elementary schools with the following letter of instructions:

> The psychological examination of more than ten thousand B-1 pupils has been completed. The report of the results of these examinations is submitted herewith.
>
> In Table I of this report, is given the percent of the B-1 pupils in the entire city that received each of the letter ratings, A, B, C +, C, C —, D or E. From this table it is seen that the 8 percent testing highest received an intelligence rating of A; the next 12 percent, B; the next 18 percent, C +; the next 24 percent, C; the next 18 percent, C —; the next 12 percent, D; and the last 8 percent, E. In other words, 20 percent of the children tested received A or B ratings, 60 percent, C +, C, or C — ratings, and 20 percent, D or E ratings.
>
> In Table II is given the percent of B-1 pupils in your own school that received each of the letter ratings. By comparing Table II with Table I, you can readily see how your B-1 pupils compare in intelligence with the B-1 pupils in all the schools combined.
>
> In Table III are given the name and letter rating of each pupil in your B-1 grade.
>
> It is recommended that you use these intelligence ratings in one or more of the following ways:
>
> 1. Arrange the pupils in three groups, the groups to be known as Groups X, Y, and Z; Group X to consist only of pupils with an intelligence rating of A or B; Group Y to consist only of pupils with an intelligence rating of C +, C, or C —; and Group Z to consist only of pupils with an intelligence rating of D or E. By this method X, Y, and Z will always designate pupils of about the same grade of ability regard-

---

[2] The revised test, copyrighted by Anna M. Engle, may be secured from the World Book Company.

less of the school in which they may be found. If for any reason, you do not find it advisable to follow the grouping suggested, you should designate the groups that you may form by symbols other than X, Y or Z. The X, Y, and Z groups in the B-1 are to be designated as I-B-X, I-B-Y, and I-B-Z. The assignment by letter ratings of pupils to Groups X, Y, and Z is to be considered tentative. You have the authority to transfer a pupil from one group to another as soon as you are convinced that he is not properly classified, provided you keep a record of the number of and the reasons for such transfers, in order that the Clinic may have definite information as to the accuracy of its classification. The Psychological Clinic stands ready to assist you in the more accurate classification of the apparent misfits by individual examination.

The pupils who are found to be able to do more work than is required of the group to which they belong are to be transferred to the next higher group, those who cannot do the work of the group to which they belong will be transferred to the next lower group, and those who cannot keep up with the lowest group, or Group Z, will be considered candidates for the special classes and will be given an individual examination by the Psychological Clinic. It is suggested in this connection, that the unassigned or coaching teachers could be of marked assistance in helping classify the apparent misfits.

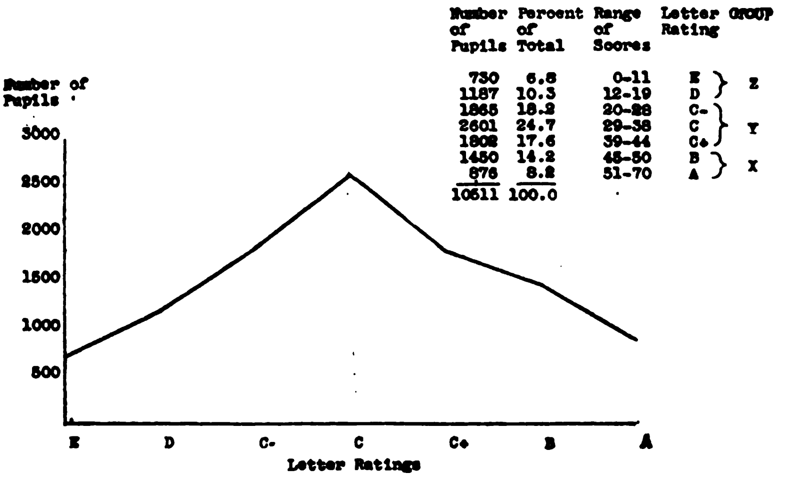
The principals who adopt this method of classification are encouraged under the direction of the Department of Educational Research to work out courses of study for the X and Z groups, an enriched course of study for the X group, and a course in minimum essentials for the Z group. In schools where there are not enough A and B or D and E pupils for a single room, it is suggested that two groups be placed in the same room, or that like groups from adjacent schools be put together.

2. In schools where all the pupils, because of congested conditions, cannot attend full time, the intelligence ratings can be used to select the pupils that stand most in need of full-time instruction. It is believed that the pupils with intelligence ratings of A, B, and C + might be so grouped and the course of study so organized that they would actually accomplish more during a half-day session than they are now able to accomplish during a full-day session in the schools where the congestion is the greatest.

3. In small schools where there are only a few cases of pupils securing an A or B rating, these pupils might be given special help by the unassigned teacher in order that they may pass quickly into the A-1 grade. This method was used to advantage in one or two of the schools last year where the first-grade children were classified by means of intelligence tests. In schools where there are only a few pupils who receive a D or E rating, the unassigned teacher could to advantage give these pupils special attention in order that the special class cases may receive an individual psychological examination as soon as possible.

The percent of pupils receiving each of the intelligence ratings deviated somewhat from the figures given in the letter of instructions to the principals. Since the points of division fell within groups making exactly the same score, rather than arbitrarily divide these groups the percents were changed somewhat. The exact percents are found in Figure 1 which gives the intelligence ratings of the 10,511 pupils in the B-1 including 1299 repeaters.

The wide range of ability found in the one hundred thirty-seven elementary schools is shown by the fact that in three schools there were no X-pupils and in seven schools there were no Z-pupils.



Number of Pupils	Percent of Total	Range of Scores	Letter Rating	GROUP
730	6.8	0-11	E	Z
1187	10.3	12-19	D	
1865	18.2	20-28	C-	Y
2601	24.7	29-38	C	
1802	17.6	39-44	C+	
1450	14.2	45-50	B	X
876	8.2	51-70	A	
10511	100.0			

FIGURE 1. THE INTELLIGENCE RATINGS OF 10,511 PUPILS IN GRADE B-1, DETROIT PUBLIC SCHOOLS, FIRST SEMESTER, 1920-21, EXAMINED IN SEPTEMBER, 1920, WITH THE DETROIT FIRST GRADE INTELLIGENCE TEST

INTELLIGENCE IN RELATION TO AGE AND SEX

The age of pupils was computed to September 1, 1920. A pupil on that date between five years ten months and six years three months was called six years of age, one between six years four months and six years nine months, six and a half years of age, and so on

TABLE I. THE INTELLIGENCE RATINGS OF DETROIT FIRST GRADE PUPILS EXAMINED IN SEPTEMBER, 1920, WITH THE DETROIT FIRST GRADE INTELLIGENCE TEST, AS RELATED TO AGE

AGES	NUMBER OF PUPILS IN EACH INTELLIGENCE GROUP														TOTALS	PERCENT OF TOTAL
	A		B		C+		C		C−		D		E			
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent		
Under 5..	1	9.1	1	9.1	2	18.2	0	0.0	3	27.2	2	18.2	2	18.2	11	0.1
5........	1	1.2	5	6.0	9	10.8	22	26.5	17	20.5	14	16.9	15	18.1	83	0.8
5½......	8	3.0	31	11.9	43	16.5	65	24.7	52	19.8	35	13.1	30	11.0	264	2.6
6........	267	6.5	501	12.2	730	17.8	1103	26.9	774	18.9	456	11.1	271	6.6	4102	39.1
6½......	243	8.1	463	15.1	540	17.6	787	25.4	520	16.9	329	10.8	186	6.1	3068	29.2
7........	131	9.2	223	15.7	232	16.3	291	20.4	258	18.1	175	12.3	114	8.0	1424	13.5
7½......	95	11.9	103	12.9	120	15.1	183	22.9	137	17.2	94	11.8	65	8.2	797	7.5
8........	47	14.7	53	16.7	63	20.0	58	18.4	38	12.0	39	12.4	17	5.8	315	3.0
8½......	28	16.1	32	18.4	22	12.6	39	22.4	25	14.4	17	9.8	11	6.3	174	1.7
9........	17	21.2	17	21.2	11	13.8	11	13.8	11	13.8	5	6.2	8	10.0	80	0.7
9½......	16	23.6	8	11.8	11	16.3	15	22.1	9	13.3	5	7.4	4	5.5	68	0.6
10 and above...	22	17.6	13	10.4	19	15.2	27	21.6	21	16.8	16	12.8	7	5.6	125	1.2
Totals..	876	8.2	1450	14.2	1802	17.6	2601	24.7	1865	18.2	1187	10.3	730	6.8	10511	100.0

Of the 9,249 pupils who entered the B-I grade in September, 1920, 3.7 percent were under six years of age; 72.5 percent were six or six and a half; 17.5 percent were seven or seven and a half; and 6.3 percent were over seven. It is evident that six and six and a half are the typical ages for entering the first grade. By reference to Table I in which are given the ages and intelligence ratings of all pupils in the B-I grade including 1,299 repeaters, it is seen that the pupils under six do not test as high as the older pupils. Apparently there is little advantage in encouraging pupils to enter the first grade before reaching six years of age, or in having them remain in the kindergarten after that age as long as provision is made for the bright pupils in the X group and for the dull ones in the Z group.

The combined distribution of the intelligence ratings of the pupils six and six and a half years of age conforms closely to the distribution of intelligence ratings for the entire B-I. In fact the distribution of the ratings of the six-year-old pupils conforms quite closely to that of the normal surface of frequency.

Of the pupils in the B-I grade 51.6 percent were boys and 48.4 percent girls. In intelligence there does not seem to be very much difference between the sexes. The girls, however, make a slightly superior showing. There were 1.8 percent more girls in the X group and 1.5 percent less in the Z group than boys. However, these differences are not great enough to have much significance.

At the end of the semester, the teachers and principals were requested to report on the ten thousand pupils who had been classified by the intelligence ratings. From this report were secured the data presented in the following tables. In order to save space, pupils are classified into the three groups, X, Y, and Z, instead of being grouped under seven headings by the individual letter ratings. It should be borne in mind that in the tables in this report and in the discussion based on the tables, groups X, Y, and Z mean the original groups formed on the basis of intelligence ratings alone—not the X, Y, Z groups at the end of the semester after many pupils had been transferred by the teachers from one group to another.

Intelligence in Relation to Promotion and Amount of Work Done

For the purpose of this study the teacher at the end of the semester was directed to report all pupils as not promoted who had not completed the work scheduled for the regular B-1 grade. The results are set forth in Table II.

TABLE II. INTELLIGENCE IN RELATION TO PROMOTION

	Group X		Group Y		Group Z		Totals	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Promoted..........	2201	96.7	5259	85.2	1168	62.6	8628	83.5
Not Promoted......	74	3.3	914	14.8	700	37.4	1688	16.5
Totals...........	2275	100.0	6173	100.0	1868	100.0	10316	100.0

It is interesting to note that more than four times as large a percent of Y pupils failed of promotion as of X pupils, and more than eleven times as large a percent of Z pupils failed of promotion as of X pupils. These facts seem to show very clearly the relation between intelligence and failure.

In the final report at the end of the semester, the teacher was asked to estimate the "amount of work done" by each pupil, taking as her standard the semester's work scheduled for the regular B-1 grade. The results are set forth in Table III.

TABLE III. INTELLIGENCE IN RELATION TO AMOUNT OF WORK DONE

Amount of Work Done (Semesters)	Group X		Group Y		Group Z		Totals	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
¼	40	1.8	429	7.4	327	20.2	796	8.3
½	58	2.7	499	8.6	300	18.5	857	8.9
¾	86	4.0	724	12.5	262	16.2	1072	11.1
1	1587	73.1	3932	67.5	710	43.8	6229	65.0
1½	199	9.2	126	2.2	7	0.4	332	3.4
2 or more	200	9.2	107	1.8	15	0.9	322	3.3
Totals...	2170	100.0	5817	100.0	1621	100.0	9608	100.0

It is surprising to discover that 28.3 percent of the pupils failed to complete a semester's work, yet only 16.5 percent failed of promotion. However, it is significant that the percent of pupils completing less than three-fourths of a semester's work is 17.2 as against the 16.5 who failed of promotion. Apparently completion of about three-fourths of a semester's work is taken as the standard for passing.

Table III shows that the number who failed to complete a semester's work is more than four times as great as the number that completed more than a semester's work; and the number that completed less than one-half a semester's work is greater than the number that completed more than a semester. Comparing the X and Z groups we find that more than six times as large a percent of Z pupils failed to complete a semester's work as of X pupils. On the other hand, 18.4 percent of the X pupils completed more than a semester's work, while only 1.3 percent of the Z pupils did so. In the X group twice as many pupils did more than a semester's work as did less than a semester's work. But in the Z group forty-two times as many did less than a semester's work as did more than a semester's work. These results seem to show clearly the folly of attempting to require all pupils to do the same amount of work in the same time. It is a marked injustice to both the dull and the bright pupil. We are urging the one to attempt the impossible and training the other in mediocrity of performance.

Intelligence in Relation to Attendance and Amount of Work Done

It may be suggested that these striking differences are not so much due to difference in ability as to difference in attendance. The number of half-days of school in the first semester of 1920-1921 was one hundred eighty-two. As a large percent of pupils attended only half-day sessions because of the congestion, each half-day session was counted as two half-days in order to make the attendance records of pupils on half-day sessions comparable to the records of those who attended full-day sessions.

TABLE IV. INTELLIGENCE IN RELATION TO ATTENDANCE AND AMOUNT OF WORK DONE

AMOUNT OF WORK DONE	GROUP X				GROUP Y				GROUP Z				TOTALS			
	Number	Less than a Semester (Percent)	One Semester (Percent)	More than a Semester (Percent)	Number	Less than a Semester (Percent)	One Semester (Percent)	More than a Semester (Percent)	Number	Less than a Semester (Percent)	One Semester (Percent)	More than a Semester (Percent)	Number	Less than a Semester (Percent)	One Semester (Percent)	More than a Semester (Percent)
Attended 150 or more half-days (82 percent or more)	1502	6.4	72.2	21.4	3793	22.6	72.1	5.3	1066	51.4	46.0	2.6	6361	23.6	67.7	8.7
Attended 100 to 149 half-days	558	11.5	75.6	12.9	1466	34.0	62.2	3.8	471	61.2	37.6	1.2	2495	34.1	60.6	5.3
Attended less than 100 half-days (less than 55 percent)	175	20.6	66.3	13.1	510	41.2	55.1	3.7	129	69.0	31.0	—	814	41.2	53.7	5.1
Totals....	2235	8.7	72.6	18.7	5769	27.1	68.1	4.8	1666	55.5	42.5	2.0	9670	27.8	64.7	7.5

About two-thirds of the pupils attended 150 half-days or more, that is, 82 percent of the time. About one-fourth of the pupils were present 100 to 149 half-days, or between 55 percent and 81 percent of the time, while 8 percent attended less than 100 half-days, or less than 55 percent of the time. There seems to be little or no relation between intelligence and attendance, as the percent of X and Z pupils attending less than 55 percent of the time is almost the same. However, the percent of Z pupils attending 82 percent or more of the time is not quite as great as that of the X and Y groups, although the difference is not great enough to have much significance. Evidently the difference between X and Z groups in the amount of work done cannot be laid to difference in attendance. In Table IV are presented the facts of attendance in relation to the amount of work done by each of the groups, X, Y, and Z.

From this table it is evident that intelligence is a much more important factor in determining the amount of work done than is a high percent of attendance. For 10 percent more failures to complete a semester's work are found among the Z pupils who attended at least 82 percent of the time than are found among the X, Y, and Z pupils combined who attended less than 55 percent of the time. Among the X pupils who attended less than 55 percent of the time is found a larger percent of pupils completing more than a semester's work and a smaller percent completing less than a semester's work than is found among the X, Y, and Z pupils combined who attended at least 82 percent of the time.

If the attendance of all pupils were increased to at least 82 percent, the percent of pupils failing to complete a semester's work would be reduced 26 in the X group, 17 in the Y group, 7 in the Z group, and 15 for all groups combined. That is, the X group would profit most, and the Z group least. Apparently most of the Z pupils are too far behind to catch up even with a much higher percent of attendance.

A similar study was made of those who failed to complete less than three-fourths of a semester's work, and it was found that if all attended at least 82 percent of the time, the number of those failing to complete three-fourths of a semester's work would be reduced 23 percent. And since, as already stated, the percent failing to complete three-fourths of a semester's work corresponds very closely to the percent not promoted, the failures would be reduced 23 percent, or 4 percent less of all the pupils would fail. Evidently if the percent of failures can be reduced almost one-fourth by increasing the attendance to at least 82 percent, the causes of absence are worthy of most careful investigation.

INTELLIGENCE IN RELATION TO SHIFTING OF PUPILS AND JUDGMENT OF TEACHER

When the intelligence ratings were sent out to the schools, each teacher was instructed to shift a pupil from one group to another as soon as she was convinced either that he could do more work than the group in which he had been placed by the intelligence ratings or that he could not do as much work as was being done by his group.

Of the 7,735 pupils reported on by the teachers, 40.6 percent were shifted out of the groups in which they had been placed by the intelligence ratings. At the beginning of the semester as classified by intelligence ratings the distribution was: Group X, 21.0 percent; Group Y, 60.7 percent; and Group Z, 18.3 percent. At the end of the semester after the teachers had shifted the pupils the distribution stood: Group X, 28.6 percent; Group Y, 47.2 percent; and Group Z, 24.2 percent. This increase in the size of groups X and Z is probably accounted for by the tendency of the teacher to fit the work in the X and Z groups to the modal child, which would make it too easy for the majority of the X pupils and too difficult for the majority of the Z pupils with the result that the best Y pupils would be better classified with the X group and the weakest Y pupils with the Z group. This explanation is further borne out by the fact that during the semester complaints were received that a large percent of the Z pupils was not able to keep up to the standard for that group; yet no complaints were received that the X pupils were not able to keep up to the standard for their group.

At the end of the semester the teacher was asked whether her estimate of the pupil's ability agreed with his intelligence rating. The results are set forth in Table V. This table brings out the fact that the intelligence ratings were confirmed by the judgment of the teacher in about 69 percent of the cases. It is worthy of remark that while 40 percent of the pupils were shifted out of the groups in which they had been placed by the intelligence ratings, yet in only about 31 percent of the cases were the intelligence ratings not confirmed by the teacher. Apparently, then almost 10 percent of the pupils were shifted, not because the intelligence rating had incorrectly measured their ability, but for other reasons. It is interesting to note that the ratings of the X pupils were confirmed in a much larger percent of cases than the ratings of the Z pupils. This is only what we should expect in view of the fact that more than one-third of the pupils come from homes where a foreign language is spoken. The judgment of the teacher, based on one semester's acquaintance with the pupil is that the intelligence ratings are correct in almost 70 percent of the cases.

TABLE V. INTELLIGENCE RATINGS CONFIRMED BY TEACHER

	GROUP X		GROUP Y		GROUP Z		TOTALS	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Ratings Confirmed..	**1696**	**77.8**	**4039**	**67.3**	**1153**	**63.4**	**6888**	**68.9**
Ratings not Confirmed...........	**484**	**22.2**	**1958**	**32.7**	**666**	**36.6**	**3108**	**31.1**
Totals...........	**2180**	**100.0**	**5997**	**100.0**	**1819**	**100.0**	**9996**	**100.0**

INTELLIGENCE IN RELATION TO BIRTHPLACE OF FATHER AND HOME LANGUAGE

Almost half of the more than ten thousand pupils in the B-1 grade are of foreign parentage on the father's side. The fathers of 25 percent of the pupils were born in Poland, Russia, or Italy. The pupils of English and Canadian parentage as shown by Table VI are equal in intelligence to those of American parentage, in marked contrast to the pupils of Italian and

Polish parentage. For less than one-third as large a percent of these pupils is found in the X group and three times as large a percent is found in the Z group as of pupils of American parentage. Yet these two groups constitute 18 percent of the total enrollment in the B-1 grade. Of course, it will be said that the reason for this marked difference is language difficulty which prevents the pupil from doing himself justice in the intelligence tests. But if this were the sole reason why do the children of Russian and German parentage test so much higher? They too, are supposed to be under the language handicap. However, it may be that a larger percent of these pupils hear English at home than is the case with the pupils of Italian or Polish parentage. Light on this phase of the subject is secured from a study of the home language.

TABLE VI. INTELLIGENCE IN RELATION TO BIRTHPLACE OF FATHER

Birthplace of Father	Group X (Percent)	Group Y (Percent)	Group Z (Percent)	Number	Percent of Total
United States......	28.7	60.3	11.0	5272	51.2
Poland............	9.0	54.4	36.6	1222	11.9
Russia............	15.4	63.0	21.6	707	6.9
Canada...........	29.6	61.2	9.2	431	4.2
Germany..........	25.2	63.2	11.6	318	3.1
Italy..............	6.1	55.9	38.0	679	6.6
England...........	29.8	62.0	8.2	218	2.1
Hungary..........	20.9	59.1	20.0	340	3.1
Other Countries....	16.1	59.1	24.8	1111	10.9
Totals..........	22.1	59.5	18.2	10298	100.0

Table VII shows that more than one-third of the ten thousand pupils in the B-1 grade come from homes where a foreign language is spoken. It is significant that the pupils who come from homes where English is spoken test so much higher than the pupils from homes where a foreign tongue is spoken. With-

TABLE VII. INTELLIGENCE IN RELATION TO HOME LANGUAGE

LANGUAGE	GROUP X (Percent)	GROUP Y (Percent)	GROUP Z (Percent)	NUMBER OF PUPILS	PERCENT OF TOTAL
English (white)....	30.0	59.6	10.4	6186	60.0
English (colored)...	23.0	63.7	13.3	443	4.3
Polish............	10.0	50.8	39.2	1258	12.2
Yiddish..........	13.2	68.9	17.9	622	6.0
Russian..........	10.7	59.6	29.7	84	0.8
German..........	16.7	64.7	18.6	193	1.8
Italian............	6.4	55.6	38.0	650	6.3
Other Languages...	15.6	56.4	28.0	897	8.6
Totals..........	23.0	59.0	18.0	10333	100.0

out question, language difficulty accounts to a great extent for this marked difference in intelligence rating. How great this language handicap is in taking an intelligence test is shown to some extent by comparing the intelligence ratings of the pupils of German parentage with those who come from the homes where German is spoken. In the former case 25 percent are found in the X group, and 12 percent in the Z group; while in the latter case 17 percent are found in the X group and 19 percent in the Z group. But while 3.1 percent of the pupils are of German parentage, only 1.8 percent come from homes where German is spoken. On the other hand, as many pupils come from homes where Polish is spoken as there are pupils of Polish parentage; and almost as many pupils come from homes where Italian is spoken as there are pupils of Italian parentage. Among other things, this may mean that the Italians and Poles acquire English with greater difficulty than do the Germans or, that they have not been in this country as long.

It is surprising to find that while 6.9 percent of the pupils are of Russian parentage only eight-tenths of one percent come from homes where Russian is spoken. However, upon investigation we found that 80 percent of these pupils came from homes where Yiddish is spoken. This evidently means that about 80 percent of the pupils of Russian parentage are Jews.

As far as I know, we have no reason to think that a child who listens to Yiddish at home has any advantage in taking an intelligence test over the child who has Italian as his home language; or, that the child who listens to German at home has any particular advantage over the child whose home language is Polish. Yet we find that twice as large a percent of the pupils who come from homes where Yiddish is spoken is found in the X group and less than one-half as large a percent in the Z group as of pupils who come from homes where Italian is spoken. And the percent of pupils in the X group who come from homes where German is spoken is one and one-half times as great, and in the Z group only one-half as great as is the case with those pupils who come from homes where Polish is spoken.

While the colored pupils test higher than any of the groups who come from homes where a foreign language is spoken, they do not test as high as the white pupils from English-speaking homes, although their average age is four-tenths of a year greater. That they apparently are rated too high as compared with the Jews, is shown by the fact that they have a smaller percent of promotions and a smaller percent completing a semester's work.

Of the groups coming from homes where a foreign language is spoken the Jews not only test the highest, but have the largest percent completing a semester's work and the largest percent of promotions. The Italians, on the other hand, test lowest, have the smallest percent of promotions, and have next to the lowest percent completing a semester's work.

While the pupils from families speaking a foreign tongue test uniformly lower than the pupils from English-speaking homes, yet, on the whole, they have been fairly well classified from the point of view of achievement.

REPORT OF PRINCIPALS ON THE X-Y-Z PLAN

In the report called for at the end of the semester, the principals of the elementary schools were requested to answer certain questions relative to the course of study for the X, Y, Z groups, the number of pupils per teacher, the promotions, the complaints of parents, and the good and bad effects of the X-Y-Z plan.

Eighty-six percent of the principals reporting expressed themselves as favoring a different course of study for each of the three groups, X, Y, and Z instead of the same course for all three groups, each group advancing at its own rate.

According to the judgment of the principals the average number of pupils that can be satisfactorily taught by one teacher in the X group is 42; in the Y group, 37; and in the Z group, 25.

Under the X-Y-Z plan, 73 principals, or 63 percent, reported more pupils completing the work scheduled for the regular B-1, 36 the "same" and 7 "less."

Eighty-seven percent of the principals reported that they had received no complaints from the parents due to the classification of pupils by intelligence ratings. No principal had any serious difficulty with parents. One case reported was that of an irate parent, whose wrath quickly subsided when informed that his child was not in the "diseased group" as reported by the child but in the "Z group."

In reply to the question as to whether the intelligence ratings had been sufficiently helpful to the principals and teachers to justify the giving of the intelligence tests, 117 or 95 percent of the principals said "Yes," four were doubtful and two said "No." Of the two who said "No," one said "No, not where the number in the B-1 is so small." There were only 17 pupils in the B-1 in that school. The other principal said "No, the teacher could gradually shift pupils according to their ability even though no test had been used."

In reply to the question asking them to state specifically any bad effects they had noticed resulting from the X-Y-Z classification, 94 or 77 percent of the principals said they had noticed no bad effects. The bad effects mentioned by the remaining 23 percent may be summarized as follows:

Some pupils misplaced by the intelligence ratings.
The Z pupils may feel stigmatized.

The teacher of the Z pupils becomes discouraged because of their slow progress.

Three groups are difficult for one teacher to handle.

The Z teacher has too many pupils to do satisfactory work.

The Y and Z groups lack incentive; competition with the more capable pupils furnished this.

Dissatisfaction of some parents who expected all pupils in the X and Y groups to be promoted.

Asked to state specifically the good effects they had noticed resulting from the classification of pupils into the X, Y, and Z groups, the principals mentioned the following:

There is more rapid progress of the brighter pupils.

The pupils' individual needs are much better met.

Poor pupils receive more help.

It has a good effect on parents who are inclined to blame the teacher because of a dull child's slow progress.

The teacher's strength and energy are better used.

It enables the teacher to adapt the course of study to suit the needs of each group.

Slow pupils do not become discouraged.

The teacher has more sympathy for the Z group, when placed in her charge as a unit.

It results in greater satisfaction to both pupils and teachers.

There is a larger percent of promotions.

No time is lost in organization at the beginning of the semester.

Teachers enjoy the work more.

There are more leaders in each group.

The slow child is aroused to do his best by proper competition.

Pupils that should be in special rooms are discovered.

It gives the teacher a much clearer vision of the work before her.

The slower pupils receive encouragement by going forward a grade a year even though they do not do as much work as the pupils in the brighter groups.

Each group is enabled to do more intensive work for time is not wasted by emphasizing work already familiar to part of the class.

It makes possible a greater study of individual differences than ever before.

SUMMARY AND CONCLUSIONS

1. By means of a group intelligence test the pupils entering the first grade were classified with sufficient accuracy to be of marked help to the teachers. The accuracy of the intelligence ratings was confirmed by the judgment of the teacher in almost

70 percent of the cases. Ninety-five percent of the principals reported that the intelligence ratings were of sufficient value to justify their being given.

2. The pupils from non-English-speaking homes tested lower than those from English-speaking homes. Language difficulty accounts in large part for this difference.

3. Of the pupils who come from non-English-speaking homes, the Germans tested the highest and the Italians the lowest.

4. Mistakes in the classification of pupils by the intelligence ratings were corrected by the teacher, who was directed to shift to the proper group any pupil who in her judgment had been misplaced by the intelligence rating.

5. The classification of pupils by means of intelligence ratings greatly increased the interest of the teacher in the study of the individual child. This fact was brought out repeatedly in the reports of the principals.

6. Most of the principals favor the development of different courses of study for the X, Y, and Z groups.

7. It is essential that different methods of instruction be developed for the bright and dull groups, that is, for the X and Z groups. Within a few weeks after the pupils had been classified by the intelligence ratings the supervisors of the elementary grades began to receive urgent calls for assistance in the teaching of the Z pupils. The teachers now realized, doubtless many of them for the first time, that they had a large number of pupils who could not satisfactorily be taught by the usual methods. In order to meet these urgent requests the supervisory staff of the elementary grades developed a little book entitled "Toys and Games," which has been of marked value in the teaching of the Z pupils.

8. The parents make no serious objection to the classification of their children by means of intelligence tests as soon as they learn that the purpose of this classification is to give the pupil a better opportunity.

9. The X-Y-Z plan is merely one step in the direction of the individualization of instruction, the equalization of opportunity.

THE FUNCTIONS OF TESTS AND SCALES AS FOUND IN RECENT EDUCATIONAL LITERATURE

A. G. Capps
School of Education, University of Missouri

The recent origin, the rapid development, and the extensive use of educational tests and scales, with indications on every hand that they will be more extensively used by all types of school people, make it fitting that we examine critically the functions which have been assigned to them. In collecting the various functions of tests and scales we examined three sets of sources: first, educational journals and magazines which are filled today with discussions of the uses or functions of tests and scales as conceived both by the school people in the field and by the research workers in our teacher-training institutions; second, recently appearing textbooks in the field of education; third, the statements of the authors who constructed the tests and scales. Most of the clerical work was done by Miss L. Alice Wilhite, a student in the School of Education, University of Missouri, and Mr. W. J. Saupe, an instructor and student in the same institution.

In our examination of these sets of sources we were amazed at the number and variety of assigned functions. One examining these claims would doubtless come to the conclusion that tests and scales form the master-key that unlocks all doors in the educational mansion. We are happy, also, to do homage to the ingenuity of educational practitioners and writers. In fact so many and so varied were the attributed functions that we found it impossible to devise a scheme of classification under which we could subsume all of them. Our task was further complicated by the fact that it was often difficult to determine, by judgment based upon the author's statements, the category to which a given function should be assigned.

In our search for a suitable scheme for classifying the functions of tests and scales we tried the following ones: (a) generic functions; (b) functions based on the school officer using the tests and scales; (c) functions based on the mental reactions of the pupils; (d) functions based on the subject matter tested; and (e) functions based on the changes in school administration, organization, supervision, and methods. Following the line of least resistance, however, we adopted Haggerty's scheme of classification[1] chiefly because his categories are fairly well defined. Nevertheless, we often found it necessary to be loose constructionists.

The following tabulation gives the results of our investigations of not less than fifty different educational books and magazine articles.

/ CLASSIFICATION OF PUPILS

	Number of Authors Mentioning
To detect the superior, mediocre, and weak students for purposes of classification and reclassification	29
To discover the special difficulties of individuals	2
To discover the psychological stages of development	1

✓ ORGANIZATION

To test one administrative device against another	4
To determine the need for organization of special classes	2
To discover definite methods of improving a program	1
To determine the optimum size of classes	1

✓ COURSE OF STUDY

To establish definite objective standards of achievement	19
To determine the subjects which need emphasis	6
To determine the standards of achievement on the basis of the relation of successive grades	4
To determine the relative value of textbooks on the basis of results secured	2
To determine minimal essentials	1
To determine the need of specialized curricula	1
To determine whether or not the course of study is properly bridging the gap between the grades	1

---

[1] Haggerty, M. E., *Specific uses of measurements in the solution of school problems.* Seventeenth Yearbook of the National Society for the Study of Education, Pt. II, 1917. pp. 25-40.

METHODS OF INSTRUCTION

To diagnose the work of individual pupils	34
To discover the efficiency of various methods	19
To stimulate invention of new devices and the improvement of methods	9
To release bright pupils from further work after attaining the standards	3
To establish in the teacher a scientific attitude	3
To determine whether or not a class is up to standard when received from another teacher	2
To demonstrate to the deficient pupil his need of drill	2
To steady the teacher's estimates	2
To detect the fact in case more time should not be spent with retarded pupils	2
To determine where to place the emphasis in instruction	1
To determine the type of drill needed	1
To enable the teachers to give help in proportion to the need of the pupils	1
To enable the teachers to give recognition in proportion to merit	1
To enable the pupils to know definitely what is expected of them	1
To stimulate the pupil to excel his own record	1

TIME DEVOTED TO SUBJECTS

To guide in the proper apportionment of time to various subjects and school activities	5
To aid in finding out how much time is required for a specific task	2

SUPERVISION

To compare classes and systems with one another and with standards	35
To measure educational products quantitatively and qualitatively	17
To aid supervisors to determine quickly and accurately the progress of students and classes	15
To reveal the actual conditions	6
To detect the strong and weak points in teacher's work	6
To diagnose the work of a class	4
To measure the progressive increase in achievement	4
To enable the superintendent and supervisors to work where most needed	3
To show the teacher that he needs better methods and when he has them	3
To give definite, unanswerable argument for the changes in a system	3

To stimulate public opinion by showing the need in definite form	3
To measure the efficiency of various features of the system	3
To obtain closer checking up of the results	2
To promote teachers on the basis of real merit	2
To enable the superintendent to locate the strong, mediocre, and weak students	2
To give definite information to parents	2
To find out the nature of the educational product	2
To test the efficacy of a new program	1
To determine the type of help and the amount	1
To furnish definite standards for judging efficiency	1
To produce steady and wholesome rivalry between the principals, teachers, and pupils	1
To enable the superintendents, principals, and teachers to see definitely the size of their task	1
To find the relation that rate bears to results achieved	1
To discover the factors that condition efficiency	1
To bring the board to see the nature of the problem	1

Since it was assumed that the foregoing functions were mutually exclusive, we found fifty-six different ones according to the statements of writers in the books and journals devoted to education.

Granting that the functions most frequently mentioned in the preceding tabulation were the most important ones, the following summarizing table was derived:

1. To compare the classes and systems with each other and with the standards.
2. To diagnose the work of individual pupils.
3. To detect the superior, mediocre, and weak students for purposes of classification and reclassification.
4. To discover the efficiency of various methods.
5. To establish definite objective standards.
6. To measure educational products quantitatively and qualitatively.
7. To aid the supervisors in determining quickly and accurately the progress of students and classes.

Functions According to the Authors of the Tests

For the most part the functions given in the preceding sections are those conceived by schoolmen in the field. Their statements indicate the uses to which they have put tests and scales in their everyday work. However, assuming that those who derived the tests and scales saw more clearly their functions, we examined statements by ten authors of tests and scales. Taking frequency of mention as a basis for deciding the relative importance of functions, we derived the following tabulation:

	Number of Authors Mentioning
To aid teaching by showing pupil status	7
To diagnose a pupil's work	5
To aid supervision by showing class status	4
To be accurate and reliable measurements	3
To aid testing methods	2
To show the differences between systems	1
To show the differences between individuals	1
To measure the ability of pupils (individuals)	1
To measure the teaching ability of teachers	1
To measure the efficiency of the text	1
To check on the teaching done	1
To locate the weakness of the pupil	1
To furnish data for prognosis and educational guidance	1
To aid in making cost studies	1
To show results in a scientific and business-like manner	1
To measure accurately some one function	1

Altogether these ten authors give sixteen different functions of tests and scales.

A comparison of the two tables of functions reveals (a) that school people make many more claims than the authors, and (b) that for the more important functions the two groups of people agree rather closely.

The more important types of purposes represented in the statements above may be summarized as:

1. Standardization—Definition of goals, subject-matter, methods, conditions, etc.
2. Classification—Pupils, teachers, etc.
3. Comparison—School systems, organizations, conditions, methods, devices, teachers, pupils, progress, etc.
4. Diagnosis—Discovery of defects and their nature.
5. Experimentation—Discovery of new and better procedures.
6. Publicity—Measurements of all kinds for use in selling campaigns.

THE THEORY OF DIFFERENTIAL EDUCATION AS APPLIED TO HANDICAPPED PUPILS IN THE ELEMENTARY GRADES[1]

J. E. Wallace Wallin
Director State Bureau of Special Education, Miami University

The theory of differential education, as well as its justification, rests upon two basic orders of demonstrated fact: first, the fact that large individual differences exist in the original and acquired make-up of children and in the rate of the growth and development of their powers and capacities; and second, the fact that children who exhibit certain well-defined kinds and degrees of specific or generic differences require special physical and hygienic treatment or educational training. This is either because the ordinary treatment does not eradicate or adequately ameliorate their handicaps, or because it does not furnish sufficient opportunity for the development of their special capacities.

Pronounced individual differences have been found in children's health, physique, and intelligence, and in their pedagogical, social, moral, emotional, and instinctive development. It is unnecessary to summarize the many studies, based on tests and observations, which have revealed these differences. Briefly stated, we know that in physical health children vary from those who score almost perfect in better babies' contests to those who border on impotence because of disease, formative defects, muscular asthenias, and dystrophies. In a detailed examination of the physicians' statements on 872 of my consecutive clinic cases, I find that the number of defects and diseases as reported in the children's past histories varied from none to 10—the latter in the case of a high grade feeble-minded girl whose diseases were chicken pox, croup, whooping cough, measles, enuresis, diphtheria, scarlet fever, pneumonia, spasms, and meningitis. The same cases varied in respect to defects revealed at the time of the examination from none to 7 (7 in the case of an imbecile girl, *viz.*, anterior obstruction or deflected septum, enlarged tonsils, right otitis media, slight deafness, dental caries, poor nervous tone, and hydrocephaly).

---

[1] Presented, in part, before the section of Psychology of the Ohio Academy of Science, April 14, 1922.

I have preferred not to discuss the question of educational differentiation as applied to superior or supernormal children. This question is more complicated than the question of training defectives, and we know far less about it than we do, for example, about the education of the deaf which has been intensively studied since 1620 (manual method) or since 1740 (oral method), and the education of the feeble-minded which has been studied since 1801.

A comparison of the scores made in anthropometric measurements by a group of 428 eight-year-old boys in a typical American city showed that those in the highest ten-percentile were 1.26 times taller than those in the lowest ten-percentile; that the best group was 3.4 times heavier than the poorest group and 3.4 times stronger, as determined by hand grip; that the best group had 2.6 times more lung capacity, as determined by the spirometer, and could perform 7.1 times more work, as determined by the ergograph. The differences found in this age group were fairly typical of the differences discovered in other age groups. X-ray photographs have shown that children of the same age may differ by six years in anatomical age. In physiological age boys may differ by five and a half years measured by pubescence, while in dentition the difference will range from 4 to 6 years according to the tooth compared.

An examination of the age-grade distribution for white children from grades I to VIII compiled as recently as June, 1920, in two large school systems which had at that time already segregated a considerable number of children in special classes, showed that 23.5 percent of the pupils were retarded from one to eight years in the one system, and 59.4 percent from one to ten years in the other (ages six and seven being considered normal for grade I). Had the defective pupils in the special classes been included the figures would have been larger. The average difference in the chronological ages of the children per grade (based on the average of the differences found between the youngest and the oldest pupils in each of the eight grades) amounted to ten years in the school system with the smaller amount of retardation. In this system there were 3.5 times as many retarded as advanced pupils. Standardized educational tests have shown that children in the same grade often differ by four or five grade steps in their educational attainments. The various factors which produce these wide variations in the pedagogical status of children of the same age are well known. One of the important factors is the difference in the ability or intelligence of the children.

In intelligence children vary from the profound idiot to the potential genius. Based on the I. Q.'s obtained by the Binet scale when given to pupils in all the elementary grades of a certain school system, the ablest pupil proved to be 2.6 times as bright as the dullest one. An average group of children of the same age in the grades have been found to differ by five years or more as determined by the same scale. Based on group tests, which we have just given to the first grade in the Miami practice school, the score of the ablest pupil was 19.5 times as high as the score of the poorest pupil in the Pressey Primer Scale, 12 times as high in the Myers Mental Measure, and 2.3 times as high in the Detroit First Grade Intelligence Test.[2] Based on the I. Q. in the Stanford Binet the ablest pupil among the first-grade children of the Miami Practice School was 1.8 times as bright as the dullest pupil (from I. Q. 77 to I. Q. 145).

We should probably find similar differences in the emotional, temperamental, instinctive, and character traits of children were it possible to apply the methods of standardized measurement to these components of personality. We know that children may vary from states of depression or dejection, bordering on hypochondria, to states of gaiety, exaltation, rapture, and euphoria, bordering on hypomania; from states of excitability, nervous erethism, and sometimes furor or "chorea insaniens," to states of profound apathy or stupor; from states of indifference, disinterestedness or emotional vacuity almost to psychopathic enthusiasm or hysterical lability of mood; from pathological reticence to uncontrollable logorrhea; from utter lack of modesty to excessive prudery; from blasphemy to fanatical religiosity; from heinous cruelty to morbid compassion.

---

[2] The large differences in the spread of ability shown in these tests is due in part at least, to the difference in the fineness of the unit of measurement. We have already shown that the spread of ability is much greater in the group than in the individual (Binet-Simon) tests of intelligence, so-called. Wallin, J. E. W., "A comparison of three methods for making the initial selection of presumptive mental defectives," *School and Society*, 13: 39. January 8, 1921.

The existence, then, of these and other differences of which we have not even hinted, constitutes the scientific presupposition on which the doctrine of differential education is based. There are doubtless many educators of distinction even to-day who, while conceding the premises, do not grant the educational consequences deduced from the facts of individual differences. They feel that the differences in large measure are not ineradicable, but that they can be leveled or attenuated by a common process of educational treatment. They also feel that the best interests of a democracy demand that at least all the children in the grades (excepting only the most extreme deviates) should be required to pursue a course of study which is uniform in its cultural and literary ingredients, and that only thus will it be possible to elevate the cultural and intellectual level of the people of the nation—a condition devoutly to be desired in a government in which the people rule.

It is not now our purpose to examine the grounds of this argument but merely to formulate briefly, and rather dogmatically, two principles of educational differentiation and segregation, which will make possible the organization of flexible, adjustable systems of special classes and schemes for providing special assistance for deficient or defective pupils in the elementary grades, and to illustrate their application to various types of such pupils. The principles may be stated thus:

First, all children who are educable but so defective physically or mentally that they are incapable of being trained for self-support, self-control, or independent existence under the condiditions obtaining in the locality in which they reside, should be assigned as early as discovered on full time to appropriate special public school classes or schools for the type of training which their disabilities require. Moreover, if the children are not properly controlled, protected, and supported at home, they should be placed under guardianship or in a residential institution, either before or after receiving training in the special classes of the public schools, and either temporarily or permanently, according to the requirements of each case.

Second, less gravely defective or deficient children who are trainable for independent existence in society but who cannot successfully cope with the requirements of the regular grades, or whose special needs cannot be adequately met by the ordinary instruction, should be assigned on full or part time to special classes, groups, divisions, or courses either for corrective training, or for special aid in the regular program of studies. Such classes, groups, or divisions should preferably be conducted in the regular grade buildings. Children thus assigned who possess the potentials for restoration or rehabilitation should be returned to the regular grades as soon as possible, while the others should be continued in the special classes or receive such special aid as may best serve their individual needs. Whenever possible two or three special classes should be organized in each center, so that the pupils may be properly graded or so that the instruction may be departmentalized.

These principles represent a defensible, conservative attitude, I believe, toward the question of special class organization. In practice they mean that the large majority of pupils in the elementary school will receive a fairly uniform training in the regular grades, while those who are handicapped will be given individual attention and specialized training as long as required.

Let us illustrate the application of these principles to the training of a few well-marked types of maladjusted school children.

Blind and deaf pupils will receive full-time assignments to special classes or schools in which appropriate methods and devices are employed and special types of training afforded. According to the 1920 census there was one blind person for every 2000 in the general population of the United States. Some estimates give the same ratio for the deaf (the 1920 census figure has not yet been announced). After a number of years of such instruction, children of good mentality might be retained in these classes only during certain periods of the day, when the teachers would prepare them for reciting with the sighted or hearing pupils in the regular grades. Pupils with only a slight degree of auditory or visual acuity, and pupils whose vision or hearing would deteriorate under use should be assigned to conservation of hearing or conservation of vision classes. The number of pupils who are thus visually handicapped has been given as 0.4 percent in a Massachusetts estimate and 0.06 percent in an Ohio estimate, the large difference probably being due to differences of standard. The number of semi-deaf, on which we have no reliable data, would vary with the standard of auditory acuity adopted. None of these pupils should be retained in these classes after medical treatment or educational training had made it possible for them to participate without undue handicap in the activities of the regular grades. Here, again, the pupils who are less handicapped physically and who are abler mentally might be prepared during certain periods of attendance in the conservation classes to recite with the normal pupils. Pupils with deteriorating vision if prepared for reciting with normal pupils should not, however, be given work which involves eye strain.

Crippled children who could not otherwise get to school should be furnished transportation. They should be kept in special schools with separate divisions for the feeble-minded and very backward, as long as they receive and require special physical treatment and educational training, such as heliotheraphy, physiotherapy, orthopedic surgery, and industrial training. When no longer in need of special physical treatment or when no further benefits accrue from such treatment, they should be transferred either to the regular grades or to such other types of special classes or schools as will adequately meet their educational and vocational needs. The government takes no census of the crippled. On the basis of a German enumeration in 1910 we should have over 133,000 crippled children in the United States.

The large majority of speech defectives can be given speech corrective training singly or in groups during certain periods of the week. The small number of speech defectives who should be segregated in special classes for more continuous training and supervision or for the protection of the other pupils, should be returned to the regular classes as soon as the disability has been corrected. It is essential in most if not in all cases to secure cooperation from the home. Based upon my St. Louis survey, almost 3 percent of all the children in the elementary and secondary schools of large cities suffer from various degrees of speech defect or disorder.[8]

Children who are ungovernable, delinquent, or truant should be transferred to special schools if feeble-minded or to ungraded class if backward. Intellectually normal delinquents and backward delinquents who do not respond to the ungraded class regime, and unstable or neurotic delinquents should be placed in special day or residential schools, separately maintained for boys and girls, and should be retained in such schools until they can be returned without detriment to themselves or to the other pupils. No accurate statistics exist on the number of pupils in the schools requiring assignment to classes of this type. But in 1918 there were 63,762 delinquent children in institutions in the United States, or six per ten thousand of the population.

Feeble-minded pupils should be segregated in special centers or schools, preferably, I believe, apart from the regular elementary school, from the time that they can be identified with reasonable accuracy until they leave school. At the age of 13 or 14 the higher-grade cases should be transferred to shop or industrial centers. I prefer separate centers for boys and girls of this type. The boundary lines for this group should be conservatively drawn both at the bottom and at the top. The regulations which I drew up for the state of Missouri provide that mental defectives should be admitted to these classes who, on the one hand, may grade as low as three years mentally or have an I. Q. as low as 30, and who, on the other hand, may grade as high as nine years mentally or have an I. Q. as high as 70, as determined by individual tests. This gives a range of seven mental years and a range of 40 I. Q.

---

[8] Wallin, J. E. W., "Report of speech defects in the St. Louis public schools." *Sixtieth annual report of the Board of Education of the city of St. Louis.* 60: 174-211, 1915-16. Wallin, J. E. W., "A census of speech defects among 89,057 public school children—a preliminary report." *School and Society,* 3: 213-16. February 5, 1916.

I am convinced, after years of experience in administering special public-school classes, that the range of mental and pedagogical ability represented between these limits is as large as it should be from the standpoint of the best interests of the children and the efficient administration of instruction, emphatically so if the pupils of these varying grades of ability must be taught in the same class. The precise limits indicated above are confessedly somewhat arbitrary, as all standards must be if they are to be rendered workable at all in the hands of different officials. Of course, one would scarcely want to enter for the first time a thirteen-year-old child with a three-year mentality in a public-school special class, nor would one want to retain very long a low-grade child who was incapable of making any improvement. Nevertheless, I have observed fairly satisfactory progress in some low-grade cases who appeared to be quite hopeless at first. One may easily become too drastic in rejecting children from the special classes as uneducable. Moreover, the benefit of the doubt should also be given to the child who grades along the upper boundary line. The upper limit of 70 I. Q. which we have set for these classes is admittedly too high in the case of some children. It is not well to be too precipitate in assigning a child to a class for the feeble-minded. It would be a very grave injustice to compel a backward child to spend all his school days in such a class. To assign to these classes children with I. Q.'s as high as 107 and to insist that they must remain, as I have found to be true, is a wanton crime against childhood.

I am quite convinced that the number of children who should be assigned to these classes in the average school system is very much less than the figures frequently suggested, namely two, three, or even four percent. For the large school system in which I spent seven years in examining educational deviates the number is only about one-half of one percent.

There is need for an intermediate class—call it "ungraded," "observation," "adjustment," or what not—for the following types of children: subnormals who are on the borderline of feeble-mindedness, having an I. Q., say, as low as 65 or possibly lower; those who are very dull or backward; those who are mentally retarded but capable of restoration; those who are of good mentality but pedagogically retarded; and those who are of uncertain mentality. The borderline cases and many of the very dull and backward would have to remain in the ungraded class until 13 or 14, when they should be transferred to an elementary industrial school. If the children are admitted while young they should be given every benefit of the doubt, and should receive a large amount of instruction in the literary branches for several years, the instruction being concretely presented. But when their limitations in the literary subject matter have been convincingly established, they should be given an increasing amount of industrial arts work, partly because of the value of the industrial training, and partly to render the literary work concrete and meaningful. These children should, of course, receive more literary instruction than the feeble-minded. Some would in time have to be transferred to the classes for the feeble-minded, while others would be restored to grade.

This borderline-backward group is many times larger than the feeble-minded group, and would in my experience, include three or four percent of the elementary enrollment in a large school system.

Above this group comes the restoration group consisting of the pupils to whom we have referred as merely mentally retarded or as merely pedagogically retarded in one or more branches. The number belonging to this group cannot be estimated with any degree of precision, but it is much larger than the combined feeble-minded borderline-backward group. This group can be reached by many types of administrative adjustment: full-time or part-time assignment to ungraded classes for intensive instruction and aid in the regular studies, with a view to restoration to a regular grade as soon as possible; assignment to small groups during certain periods of the day for coaching in weak branches; reciting with a lower grade, as well as with their own grade, in the branches in which the pupils are weak; the provision of assistant teachers according to the Batavia plan to bring up the laggards and to insure equal progress for all; the organization of parallel differentiated courses of study, according to the Cambridge plan,[4] or the Baltimore (1898) and Santa Barbara plans;[5] the institution of quarterly or semi-annual systems of promotion; the improvement and motivation of the instruction in the regular grades, especially by rendering it more concrete and meaningful (through the use of correlation, cores of interest, projects, etc.).

In each of the above groups except the lower grades of the feeble-minded a small number of pupils are subject to special pedagogical disabilities, resting upon a central defect, which seem to be largely ineradicable. Thus, some pupils varying in general intelligence from the high-grade feeble-minded to the supernormal never acquire any competency in arithmetic or in reading. I am more familiar with the latter defect, having found among 2,774 deficient children coming to a psycho-educational clinic 98 children classifiable as cases of dyslexia, and 16 as cases of visual aphasia; or a total of 114 word-blind cases, 95 of whom have been subjected to special analysis.[6]

Finally, how shall we discover the children in need of special consideration? We should have available for some or for all of the pupils the results of thorough physical examinations, of inquiries into the developmental, pedagogical, personal, and family histories, and of standardized educational and psychological tests. Over a decade of experience in making more than 3,500 clinical examinations has convinced me that where it is possible to secure a carefully filled out pedagogical record from a well-trained, intelligent teacher, the results of the so-called attainment or efficiency tests are not essential for a satisfactory evaluation of the child's pedagogical status. Trained and experienced teachers and principals can, if they will, rate pupils in their school work more accurately than they are given credit for in the current test literature. They can offer pedagogical information regarding a child which it is absolutely impossible to secure from any standardized educational test. Nevertheless, even in such cases the attainment tests furnish corroborative evidence or a check on the school rating. Naturally when no school reports are available or when only indifferent or poor reports can be obtained, the scores from the standardized tests are of very considerable value.

---

[4] The revised, 1910, plan provides two parallel courses, the basal course (for most pupils) requiring eight years and the other six years (for gifted pupils) to complete. The gifted pupils do one-third more work each year. A number of transfer points between these courses make it possible for pupils to take any amount of time from six to eight years or even longer, to complete the grades.

[5] During the first six years there are three courses, namely, for the slow, average, and superior pupils. All the pupils spend the same amount of time in each grade, but the abler pupils secure a richer course. The slowest pupils cover only the minimum essentials. Pupils can be transferred from group to group according to individual requirements. Above the sixth grade the work is departmentalized, many studies are provided, and promotion is by subjects. The superior students are permitted to take extra subjects.

[6] Wallin, J. E. W., "Congenital word-blindness—some analysis of cases." *Training School Bulletin*, 17: 76-84, 93-99. September-October, 1920.

Recently we have entered upon a period of trenchant criticism of the assumptions, validity, value, and claims of the group tests of intelligence. This wave of criticism is partly due to the extravagant claims made by the test publishers, and by an avalanche of uncritical and laudatory articles put out by a legion of amateur scientists. It is also due to the lack of consistency among the scales themselves and to the discrepancies frequently found between the test results and other criteria. When the doctors disagree the patients suffer, and the public becomes alarmed. When different group tests disagree in the rating of any considerable number of pupils the validity of the tests is immediately brought into question. In comparing the ratings obtained in five group tests of intelligence given to a class selected at random from the upper elementary grades of a school in New York, it was found that for over 18 percent of the children the difference in the mental age ratings varied from 25 months to 31 months as compared with the average rating obtained in all five tests. The corresponding differences when the individual scales were compared, having regard for the plus and minus deviations, varied from 38 to 49 months (Stenquist).

Recently at the Miami University practice school the pupils in the sixth, seventh, and eighth grades were given the Stanford-Binet, the National, the Illinois, and the Pintner Non-Language group tests. The pupils in the first grade of the same school were given the Stanford-Binet, the Pressey Primer, the Myers Mental Measure, and the Detroit First Grade Intelligence Test. The median difference in the intelligence age rating of the upper-grade pupils obtained by the Stanford-Binet and the Illinois Examination (the only scales with directly comparable norms) amounted to 11 months, varying from no difference to 46 months. The differences ranged from 13 months to 46 months in 43.8 percent of the cases, and from 25 to 46 months in 17 percent.

Among the first-grade pupils the discrepancy was quite as great. The material is being analyzed and will be made available in another article. On the basis of ranks the displacement of individual pupils was from 0.5 to 31.5 steps. This is to be taken in connection with the fact that the number of pupils taking all the tests was 34.

If for each test the pupils are placed in quartiles no two tests agree on more than a third of the pupils who should be assigned to either of the middle quartiles. Even in respect to the first and fourth quartiles, where agreement would be expected to be greatest, in only four out of ten comparable pairs of ratings do any two tests agree on more than half of the pupils who should be assigned to these quartiles.

Psychologists have for years been proceeding on the assumption that a fairly high correlation coefficient obtained by one of the recognized correlation formulas is proof of the value or validity of the test. But the correlation coefficient has often, in my experience, proved to be a clever device for concealing the true state of affairs—a fictitious refuge in our search for security. With a high correlation coefficient between a group test of intelligence and the Binet-Simon individual test, I have frequently discovered large discrepancies in rating when a careful analysis has been made of the individual cases.[7] Plus and minus deviations are not neutralized or concealed when individual cases are compared. After all, differential education must rest upon a careful consideration of the special needs of particular individuals; and we must know whether each of those with whom we are concerned are accurately scaled by the test used. Even a high correlation coefficient can not demonstrate the accuracy of a particular measurement. Moreover, the use of group intelligence tests as a means of prognosticating success in the different subjects of the curriculum has been beset with grave difficulties because of the low correlations frequently found, especially among college students. The correlation coefficients have varied from almost zero to 0.60 or 0.70, most of them being less than 0.50. In the junior and senior high schools the correlations seem to be higher, but they very rarely reach the degree of correlation required by Toops in tests of vocational aptitude. He is reported to require a correlation of over 0.75 or 0.80 with actual achievement before the tests are of value to the employer, or before they have any advantage over present methods of selection. Most of the literature regarding the success of sectioning pupils in the elementary grades on the basis of group intelligence tests is extremely laudatory, but it is sadly in need of critical evaluation.

I have never ventured to make any individual diagnoses on the basis of the results from group intelligence tests. I have felt that they might be of considerable value in making preliminary selections of the most deficient children for further examination. But, in an experiment with a group intelligence test in the St. Louis schools, I found that, by selecting the pupils reported as the most deficient in intelligence by the teachers and principals on a prepared form in accordance with prescribed instructions, a slightly larger percent of mental defectives was secured than by selecting the pupils who made the lowest score in the group test. Some of those who made a very low score or no score proved to be normal or nearly normal when individually examined.[8] It should be remembered, however, that those who were reporting my cases had had considerable experience in proposing children for examination as candidates for classes for the feeble-minded. In another school system the group tests might have proved superior to the school staff.

---

[7] Wallin, J. E. W., "A comparison of three methods for making the initial selection of presumptive mental defectives." *School and Society,* 13: 31-45, January 8, 1921. Other data are in preparation.

The schools will probably continue to employ group intelligence tests as an aid in determining who among the extreme deviates shall be selected for an individual examination. This will be especially true where there is no other available method of comparable reliability. But it may be doubted whether the practice of running all children through group intelligence tests as a periodic, routine measure will justify the time and expense involved; or whether we can group children with accuracy into fast, average, and slow sections purely on the basis of group intelligence tests, certainly not until the tests measure more consistently than is the case at present. Even then, many other factors must be considered in educational grouping, such as special interests, abilities and disabilities, physical and emotional maturity, health, and prior training. We are, of course, speaking of the group tests in current use, all of which measure imperfectly only a small segment of intelligence, or of intelligence plus attainment under arbitrary and unnatural conditions. Some time in the future with more highly perfected group tests it may be possible to do the things of which I am now dubious.[9]

With respect to the use of the individual scales, such as the Binet-Simon, I think there can be no doubt of their value, especially when they are applied to the type of individuals for whom they were devised, namely, the feeble-minded and seriously deficient children in the elementary schools. Over twelve years of constant employment of the Binet-Simon scale in clinical examinations has convinced me that it has considerable value as a measure of the potential ability of the feeble-minded and very inferior for wrestling with the work of the schools. This is no doubt due in part to the fact that the scale is to some extent an educational scale (It certainly is not a test of "pure native ability," whatever that may mean). Although still imperfectly constructed and standardized, this scale gives us an independent estimate of a child's ability, which is practically indispensable in the absence of authenticated histories or other reliable evidence. When reliable facts are otherwise available on which to base a judgment, it adds additional certainty to the diagnosis; and in all cases it enables us to get a more accurate rating of a type of ability which we have been designating general intelligence than can be otherwise secured. Certainly the Binet rating, even allowing for a margin of error, is incomparably superior with the vast majority of subnormals to any intelligence rating or mental diagnosis based on a physical examination. It is notoriously true that a physical examination gives no accurate knowledge of mental status except in a few special types in which the associated mental peculiarities are well known. The Binet tests, or a similar scheme of tests, will probably always render an indispensable service when used in conjunction with tests of motor, mechanical, and educational ability, and with the discretion and insight which come from a knowledge of anomalies of behavior in the fields of volition, motility, intelligence, instinct, emotion, and character.

---

[8] *Ibid*, p. 31 ff.

[9] One cannot, however, even with accurate measures classify pupils for instruction purely on the basis of the intelligence age or the I. Q.

Such tests are not without value in the hands of anyone who can give them accurately and who understands the art of making individual psychological examinations, but they are not a short-cut to erudition or a substitute for the scholarship and interpretative power which come from years of specialized study of the abnormal cases themselves. In using intelligence tests, whether group or individual, let us not forget that the securing of a test score is the *beginning* and not the *end* of the process. The securing of a score has all too frequently been considered the "be-all and the end-all" of an examination. A test score is something calling for further analysis. It is an x to be interpreted. To suppose that any intelligent person who can administer tests is competent adequately to analyze, interpret, and classify and to serve as director of a psycho-educational clinic or of a so-called bureau of tests and measurements concerned with psycho-educational diagnoses is to fly full in the face of the demonstrated complexity of human behavior and classification.[10]

---

[10] The writer has consistently challenged this assumption during the past decade (see *The Mental Health of the School Child* and the *Problems of Subnormality*). The American Psychological Association has by a resolution called attention to the qualifications needed by psychological examiners, the clinical section of this association has established high eligibility qualifications for membership, while the State of New York has established minimum requirements for psychological practitioners. In spite of these facts two amateurs were last year appointed at the salaries of university professors as heads of divisions in two large public school systems having charge of the classification of mentally abnormal children. One of these persons (who is receiving the highest salary ever paid in the position) had never before given a single educational or psychological test, nor made any study whatever of the pathology, symptomatology, or nosology of mental abnormalities. The other had just finished an arts course in college. When a prominent social worker in one of these cities protested against this kind of appointment, he was told that an expert was not wanted in that system, but a person who could give tests, take orders, and leave it for some one else to classify and assign the children as he saw fit. Such practices probably will not be curbed until a constituted committee of the American Psychological Association turns the light of publicity upon the offending systems.

THE CALCULATION AND INTERPRETATION OF PERCENTILE RANKS

L. L. Thurstone
Carnegie Institute of Technology

ABSOLUTE AND RELATIVE RANKS

It is sometimes desirable to state the relative standing of an individual with reference to the other members of a group. This is usually done by giving him a rank. Thus if there are seventeen persons in a group the person who ranks highest is given a rank of seventeen and the lowest person is given a rank of one.

If we are told that a person has a rank of 27 we do not know whether he ranks high, average or low, unless we also know the number of individuals in the group. If a person ranks 27 in a group of 29, he ranks high relative to the group. If he ranks 27 in a group of 1000 he is relatively low. When designating an individual by his *absolute* rank one must also state the number of individuals in the group with which he is compared. The absolute rank of an individual member of a group is his numerical position when all the members of the group have been arranged in order from the lowest to the highest. The highest numerical rank is given to the person who ranks highest. In this sense the designation of absolute ranks differs from the customary expression of calling the best person "first." In statistical language "one" means the lowest numerical score, no matter what the meaning of the score may be.

In order to avoid the necessity of stating the number of cases involved in designating a person's relative position by rank one may express the absolute rank in terms of *percentile rank*. In this case one states the rank that the person would have if there were one hundred members in the group. If there are fifty individuals in a group the middle person would have an absolute rank of 25 but his percentile rank would be 50. A percentile rank is so calculated that it indicates the percent of the group which ranks below the specified percentile. A person who has a percentile rank of 72 exceeds 72 percent of the group and is exceeded by 28 percent of the group.

The median score is always the 50th-percentile. The upper quartile is always the 75th-percentile, and the lower quartile is always the 25th-percentile.

When several individuals have the same score, a certain assumption is made concerning their ranks. Consider the ten individuals designated by the following scores:

Individual:	A	B	C	D	E	F	G	H	I	J
Score:	16	24	24	35	41	56	56	56	72	83

These scores are arranged from the lowest to the highest. Now if we designate these indivduals and their scores by absolute ranks, we have the following arrangement:

Individual:	A	B	C	D	E	F	G	H	I	J
Score:	16	24	24	35	41	56	56	56	72	83
Absolute Rank:	1	2	3	4	5	6	7	8	9	10

But B and C have the same score and consequently they should have the same rank in order to state fairly their relative position in the group with which they are compared. This is done by assigning to each of these two persons the same rank, namely their median rank. The revised statement of absolute ranks is then as follows:

Individual:	A	B	C	D	E	F	G	H	I	J
Score:	16	24	24	35	41	56	56	56	72	83
Absolute Rank:	1	2½	2½	4	5	7	7	7	9	10

In this case B and C who have the same score, 24, are given the same absolute rank, namely 2½, and the three individuals, F, G, and H who have the same score, 56, are given the same absolute rank, namely 7. The total number of ranks assigned is ten because there are ten individuals in the group. No individual is given the rank of 6 or 8, these being skipped in order to balance the ranks for the small groups of equal scores.

The same assumption is made concerning percentile ranks. Suppose that one hundred individuals are ranked from the lowest to the highest and that the twenty highest individuals all have the same score. It would be unfair to assign a percentile of 81 to one of these twenty persons and a percentile of 100 to another when both have the highest possible score. In cases of this sort all twenty are given the same percentile rank, namely 90, which is the median precentile rank of the twenty individuals who have the same score at the top of the group. The same assumption is made for groups of identical scores in the middle range of the distribution and at the lower end. As a consequence it is possible to assign percentile ranks in such a way that no individual gets a percentile of 100 even though several of them have the highest possible score.

CALCULATION OF PERCENTILE RANKS

In the accompanying table we have the calculation of the percentile ranks for the Liberal Arts freshmen at the University of Texas in an intelligence test. The first two columns indicate the class intervals, two columns being used to specify the lower

TABLE I. SCORES AND PERCENTILE RANKS, TEST IV, LIBERAL ARTS FRESHMEN, UNIVERSITY OF TEXAS

SCORE			FREQUENCY	PERCENTILES		
From	To	Mid		Lower	Upper	Mid
1	*2*	*3*	*4*	*5*	*6*	*7*
0	9	5	0	....	....	....
10	19	15	0	....	....	....
20	29	25	6	.000	.006	.003
30	39	35	19	.006	.029	.017
40	49	45	56	.029	.094	.061
50	59	55	94	.094	.203	.148
60	69	65	154	.203	.382	.292
70	79	75	161	.382	.569	.475
80	89	85	143	.569	.735	.652
90	99	95	95	.735	.845	.790
100	109	105	65	.845	.921	.883
110	119	115	33	.921	.959	.940
120	129	125	16	.959	.978	.968
130	139	135	8	.978	.987	.982
140	149	145	7	.987	.995	.991
150	159	155	3	.995	1.000	.997
160	169	165	0	....	....	....
	Total............		860	....	....	....

Rate = $1/n = 1/860 = 0.001162$

and upper limits of each class interval. The third column gives the midpoint of the class interval. The fourth column gives the frequencies. The sum of the frequency column is of course the total number of cases, *n*, which in this illustration happens to be 860.

The next step is to label the three percentile columns, the lower percentile, the upper percentile, and the midpercentile, respectively. This is done because the percentile rank which belongs to the bottom of a class interval is not the same as that which belongs to the upper edge of that class interval. The midpercentile between these two is assigned as the percentile rank for the whole class interval. The failure to take this bit of logic into consideration accounts for many inconsistencies in the calculation of percentile ranks.

The first step in the actual calculation is to determine the *rate* which is the reciprocal of the total number of cases. In the illustration of the accompanying example, the total number of cases is 860 which is *n*. The rate is $1/n$ or 1/860 which is .001162. This can be determined either by performing the indicated division or directly from Barlow's Tables or any other similar tables.

Every class interval represents a percentile range. In the fifth column we have the lower limit and in the sixth column the upper limit of the percentile range for each class interval. The upper limit of the percentile range for any given class interval is identical with the lower limit of the percentile range of the next higher class interval. This may be seen by inspecting the tabulated calculations. The percentile rank which is finally given to all the individuals in any class interval is the midpoint of the percentile range for that class interval. This is tabulated in the last column.

After the rate has been determined one registers this number in a calculating machine. In the illustration here given the rate would first be added six times (corresponding to the first frequency of 6); and the machine would then show .006 which is the upper percentile rank for the class interval 20-29. The lower percentile rank for the lowest class interval is of course zero and the upper percentile rank for the highest class interval is necessarily 1.00. Add the rate nineteen times more in the machine and it will then show .029. This is in reality nothing but cumulative addition. The upper limit .006 for the class interval 20-29 is identical with the lower limit .006 for the class interval 30-39. Record these percentiles and add the rate cumulatively 56 times which will show .094 as the upper percentile for the next class interval.

When the whole column has been completed, one will have added the rate *n* times (.001162 × 860) which will give 1.00 for the upper limit of the highest class interval. The calculation of percentile ranks is greatly facilitated by using a calculating machine such as the Burroughs adding machine with the attachment by which the setting may be shifted to the right and to the left.

Stated briefly the percentile ranks are calculated by adding the rate cumulatively as indicated by the frequency column. The sum thus obtained immediately before adding the frequency for any given class interval is the lower percentile rank for that class interval. The sum obtained after the class frequency times the rate has been added is the upper percentile rank for that class interval. With the help of an adding machine the calculation of percentile ranks such as those of the illustration requires only three or four minutes. Any clerk can be taught to do this work; and the method here described affords an automatic check on the accuracy, because the last sum must necessarily be unity.

THE PERCENTILE CURVE

The percentile curve is a curve which shows the relation between the given variable, such as scores, and percentile ranks. By means of a percentile curve one may ascertain at a glance the percentile rank corresponding to any given score. The rank is of course with reference to the group whose distribution is represented by the curve. The percentile curve is in reality a cumulative frequency curve in which the frequency ordinates have been stated as fractions of *n*.

In Figure 1 we have the percentile curve for the frequency table and percentile calculations in the accompanying table. The base line represents scores in the test and the ordinates represent percentile ranks. By consulting the table we find, for example, that the class interval 80-89 has a percentile rank of .65. This agrees with the curve which has an elevation of .65 for the mid-point of the class interval 80-89.

Note that the class interval 80-89 has a corresponding percentile range from .569 to .735. This range is indicated on the graph by a short vertical line extending from the level .569 to the level .735. We can therefore read directly from the chart the percentile range for each class interval. Note that this vertical line is drawn not at either end of the class interval but directly above its mid-point. The curve is plotted by drawing these vertical lines as indicated by columns 5 and 6 in the table. The curve is drawn through the mid-points of these vertical lines. In practice one may therefore eliminate the calculation of the mid-percentiles since these can be plotted by inspection with sufficient accuracy for all practical purposes. When one becomes accustomed to this work it is possible to eliminate the third, sixth, and seventh columns in the table and to determine these facts graphically.

In practice it is found that with a small number of cases the curve would not be smooth if it were drawn strictly through the mid-points of the vertical lines. For this reason one sometimes generalizes the curve by smoothing it so as to pass as nearly as possible through the mid-points of the successive vertical lines. This gives a *smoothed percentile curve*. The reasoning by which a percentile polygon is smoothed into a percentile curve is similar to that of the smoothing of a frequency polygon into a smoothed frequency curve. However, when a frequency curve has been smoothed it is only fair to the reader of one's report that the vertical lines be inserted because they represent the actual observations. The reader can ascertain at a glance to what extent the curve has been smoothed by noting the deviation of the curve from the mid-points of the vertical lines. It should also be observed that the percentile polygon is more readily seen to be continuous than the corresponding frequency polygon. Even when the frequency polygon is quite irregular, the percentile polygon or curve which is plotted on the same drawing and to the same scale will be relatively continuous or smooth.

PROPERTIES OF THE PERCENTILE CURVE

(1) The percentile curve is asymptotic to the base line or zero and also to the percentile rank of 100. The percentile curve theoretically never reaches zero nor 100. Therefore it is theoretically impossible for any individual to have a percentile rank of zero or a percentile rank of 100. This fact is probably not generally recognized.

(2) The median score can be read directly from a percentile curve because it is the score which corresponds to the 50th-percentile. In the illustration we find that the 50th-percentile gives a score of 77 which is the median of the group. This saves considerable labor in calculating the balanced median.

(3) The upper quartile can be determined in a similar manner because the upper quartile point is the 75th percentile. In the illustration the 75th-percentile corresponds to a score of 92. The lower quartile point corresponds to the 25th-percentile which in the illustration is 63. When the percentile curve has been plotted it is therefore not necessary to calculate the quartile points because these can be read directly from the graph. From these facts one may of course also determine graphically the quartile range or the semi-interquartile range as desired.

(4) The mode is defined as the point on the base line of a frequency curve directly under the highest ordinate of the curve. The mode is therefore that score which occurs most frequently in the distribution. This can be determined approximately from the percentile curve by selecting the point on the base line directly under the steepest part of the percentile curve. Since the percentile curve is in reality a cumulative frequency curve it should be clear that the steepest part of the percentile curve is directly over that score which occurs with the highest frequency. While this graphical determination is only approximate it may

be sufficiently accurate for practical purposes and it saves considerable labor in computation.

(5) When two distributions are to be compared, considerable information may be obtained by simply inspecting the percentile curves for the two distributions. Thus, for example, the ranges of the distributions may be compared by inspecting the respective ranges of the two percentile curves. The variabilities of the two distributions may be compared by inspecting the relative slopes of the two percentile curves. The steeper the general slope of the percentile curve the smaller is the variability of the corresponding distribution. The two medians may also be compared graphically.

CONSTRUCTION OF THE PERCENTILE CURVE

When a percentile curve is to be constructed it is not necessary to perform all of the calculations indicated in the accompanying table of data. I prepared that table so that it would be complete and so that it would show all the steps that are implied in the calculations.

In practice it is only necessary to calculate the figures in columns one, two, four, and five. But when the calculations are made with an abbreviated form it is absolutely necessary to keep in mind that the percentiles recorded in the fifth column are not the percentiles that are to be assigned to the adjacent class intervals. The percentiles recorded in the fifth column indicate the lower percentile of the percentile range for each class interval. In order to locate the upper percentile for any class interval it is necessary to consult the percentile recorded in column five for the next higher class interval. This becomes easy with practice because the figures are then immediately transferred to a chart which facilitates the interpretation.

GRAPHICAL CALCULATIONS OF PERCENTILES

It is possible to calculate percentile ranks without any numerical calculation whatever. I have illustrated a graphical method in Figure 2. Let us suppose that we have a frequency table with the scores arranged in class intervals and the corresponding class frequencies. Let there be 160 cases in the distribution arranged in class intervals of ten along the range from zero to 170. Plot the accumulated frequencies in the form of vertical lines over the mid-points of their respective class intervals. The lengths of these lines would be equal to the actual frequencies according to the frequency scale at the right end of the diagram. Draw a smooth curve A through the mid-points of these vertical lines. This is the accumulated frequency curve. In the diagram the lines representing the class frequencies have been omitted in order to avoid confusion. The accumulated frequency curve A can be used to determine the absolute rank of any member of the group. What we want, however, is a similar curve which gives the percentile rank instead of the absolute rank. Draw on the left edge of the diagram any suitable scale running from zero to 100 to represent percentile ranks.

The graphical construction is made as follows. Locate any point C at random on the elevation of the top of the curve A. Locate the point E directly under the point C and on the elevation of 100 on the percentile scale at the left of the diagram. From these two points, C and E, draw straight lines to any point D on the base line of the diagram.

Assume any point F on the curve A. Draw a horizontal line from this point to its intersection with the upper diagonal line at G. Drop a vertical line from G to its intersection with the lower diagonal line at H. Draw a horizontal line from H to a point K directly under the starting point F. This point K is the point on the percentile curve corresponding to the point F on the curve of absolute ranks. Proceed similarly with a number of other points on the curve A. In the lower part of the diagram I have indicated the same procedure for another point F′ on the curve A.

Draw a smooth curve through the points so located. This will be the desired percentile curve B. The procedure is very easily carried out on a drawing board, in fact, much easier than the verbal description would indicate. In plotting the original curve A of accumulated frequencies one should be careful to draw the curve through the mid-points of the vertical lines which represent the class frequencies.

An automatic check on the plotting of this curve is obtained in the fact that the total elevation of the vertical lines for curve A should equal the total number of cases in the distribution which in the illustration was 156.

Where suitable calculating machines are available I should not recommend this graphical procedure. It is then best to deal with the numerical frequencies and the numerical forms of the percentile ranks as indicated in the accompanying frequency table.

It is essential that those who use percentile ranks considerably in the tabulation of educational and psychological measurements should understand the fundamental properties of the percentile curve. If one becomes accustomed to the calculation of the percentile ranks correctly, one will not be inclined to tolerate rough and ready guesses for the percentile ranks that should be assigned to given scores. The most frequent error in the assignment of percentile ranks to given scores is in the form of a con-

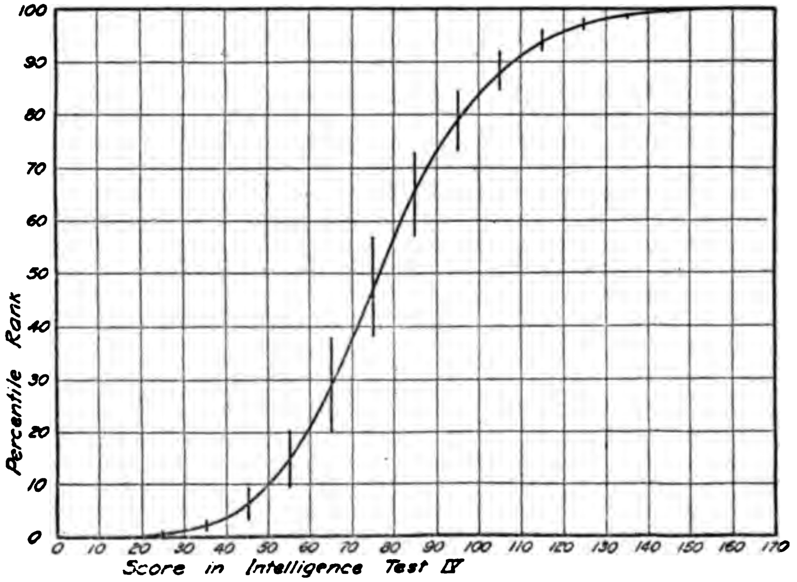


FIGURE 1. PERCENTILE CURVE CORRESPONDING TO TABLE I.

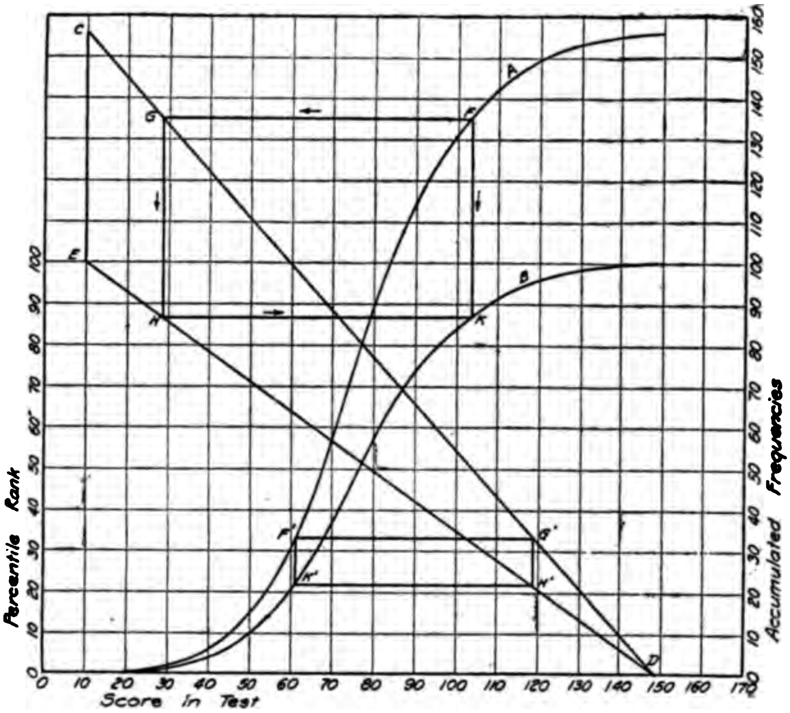


FIGURE 2. A GRAPHICAL METHOD OF CALCULATING PERCENTILE RANKS

stant error. This is caused by the fact that the percentile is determined either for the upper edge of a class interval or for the lower edge. This percentile is used for the whole class interval instead of the one which corresponds closely to the mid-point of the interval. The error is avoided by following the procedure illustrated in the accompanying frequency table with percentile calculations.

A CRITICISM OF TESTS REQUIRING ALTERNATIVE RESPONSES

H. H. HAHN
State Normal School, Wayne, Nebraska

Before further use is made of "Yes-No Questions," "Same-Opposite Tests," "Alternative Questions," or "True-False Statements" for testing mental ability or achievement in subject matter, the nature and reliability of such tests should once more be critically examined. In this article this form of tests will be discussed, first, as to its accuracy and, second, as to its influence on teaching when used to measure achievement.

Obviously one of its dangers is the guessing factor. Because the right answer is either "Yes" or "No," "True" or "False," "Same" or "Opposite," one has a chance of *guessing* one-half of the answers correctly. For this reason an attempt must be made in scoring to eliminate the effect of guessing. The method adopted is to subtract the wrong from the right answers, and to accept the remainder as an accurate measure of ability or achievement. But does *subtracting the wrong answers from the right* eliminate the effect of guessing from the test?

If one holds that by subtracting the wrong from the right answers one eliminates the guessing factor, no more and no less, one must assume that individuals in such tests *always guess an even number of times;* for if they should guess an odd number of times the effect of guessing could not be wholly eliminated by this method. One must further assume that if an individual happens to guess wrong the first time *his second guess must be right, his third wrong, his fourth right,* and so on, guessing wrong and right alternately; for if he should guess wrong or right twice in succession, the method could not eliminate the guessing effect. That the law of chance does not operate even approximately in this way can easily be demonstrated.

For the following experiment seventy buttons were used. These buttons were alike in every way except that thirty-five were white and thirty-five were black. They were placed in a small paper sack and thoroughly mixed. Then, a blindfolded person drew out twenty-five buttons, one by one, the sack being shaken well after each drawing. This was repeated 25 times with the results shown in Table I. In this experiment the 25 repetitions of the performance are to represent 25 individuals taking a yes-no test. The white and black buttons indicate wrong and right guesses, respectively. Therefore, in Table I *w* means white button or wrong guess, and *b* means black button or right guess. The capital letters at the left of the table indicate the 25 performances, and the numbers at the top indicate the number of drawings. If we assume now that each individual guessed five times in answering the 25 questions (the first five answers being the guesses), and if we score the results by the method of eliminating the guessing factor, we find that the respective scores of the individuals tested are as follows: 0, 15, 0, 0, 5, 0, 5, 0, 5, 0, 5, 0, 0, 0, 25, 15, 0, 15, 5, 5, 5, 0, 5, 15, 5, 5, 5.

TABLE I.—RESULTS OF DRAWING 25 BUTTONS 25 TIMES FROM A BAG CONTAINING 35 WHITE AND 35 BLACK BUTTONS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
A.....	w	b	w	b	w	b	w	b	b	b	b	b	b	b	b	b	w	b	b	b	w	w	b	b	w
B.....	b	b	b	w	b	b	w	w	b	w	w	w	b	w	b	b	w	w	b	b	w	w	w	b	b
C.....	w	w	b	w	b	b	b	b	w	w	w	b	w	w	w	w	b	w	w	b	b	b	b	b	w
D.....	w	w	w	w	w	b	w	b	b	b	b	b	w	w	w	w	w	w	b	b	w	w	w	b	b
E.....	b	b	w	w	b	w	w	w	w	w	w	w	w	b	b	w	w	w	w	w	w	b	b	b	b
F.....	w	w	b	w	b	b	b	w	w	w	b	w	b	w	w	w	b	w	b	w	w	w	b	b	b
G.....	b	w	b	b	w	b	b	w	w	w	w	b	b	w	w	w	w	w	b	w	b	w	w	b	w
H.....	w	w	w	w	b	w	b	b	w	b	w	w	w	b	w	w	b	w	w	w	w	b	w	b	w
I......	w	b	w	b	b	b	w	w	b	w	w	w	b	b	w	b	b	w	w	b	b	w	w	w	w
J.....	w	b	w	w	b	w	b	w	w	b	b	w	b	b	b	b	b	b	w	w	b	w	b	b	b
K.....	w	w	w	w	b	b	b	b	b	b	w	b	w	w	b	w	b	w	b	b	b	w	w	w	b
L.....	b	w	w	w	b	b	b	w	w	b	w	w	w	w	b	b	b	b	b	b	w	b	w	b	w
M....	b	b	b	b	b	w	b	b	b	w	b	b	b	b	w	w	w	b	w	b	b	b	w	b	w
N.....	b	b	b	b	w	b	b	b	w	w	b	w	b	b	b	w	w	b	b	w	w	w	w	w	w
O.....	w	w	w	b	b	w	b	b	b	b	b	b	b	w	b	b	w	b	w	w	b	b	w	b	w
P.....	w	b	b	b	b	b	w	w	w	b	w	b	w	b	b	b	w	w	w	b	w	w	w	b	w
Q.....	w	b	w	b	b	w	b	b	b	b	b	w	w	b	w	b	w	b	w	w	w	b	w	w	b
R.....	w	b	w	b	b	w	b	b	w	w	w	b	w	b	w	w	w	w	b	b	b	b	b	b	b
S.....	b	w	b	w	b	b	b	b	b	w	b	w	b	w	b	w	w	w	b	b	b	b	w	w	w
T.....	b	w	b	w	w	b	b	b	b	w	b	b	b	w	w	w	b	w	b	b	w	w	w	b	w
U.....	b	w	w	b	b	b	b	w	w	w	w	b	b	w	b	b	w	b	w	w	b	w	w	w	b
V.....	w	b	b	b	b	b	w	b	b	w	b	b	w	w	b	w	b	w	w	w	b	w	b	w	b
W....	w	b	b	b	w	w	w	w	w	w	w	b	b	w	w	b	b	b	b	b	w	w	b	b	b
X.....	b	b	b	w	w	b	b	b	b	b	b	w	b	b	w	b	w	b	b	b	w	w	w	b	w
Y.....	b	w	b	w	b	w	w	w	b	b	b	b	w	b	w	w	b	b	b	w	b	w	w	b	w

The median of the whole group is increased by 5 and the average by 6. If each guesses seven times, their respective scores are increased by 5, 15, 5, 0, 0, 5, 15, 0, 5, 0, 0, 5, 25, 25, 0, 15, 5, 5, 15, 5, 15, 15, 0, 15. In this case the median is increased by 5 and the average by 8. If every answer is a guess, then their respective scores are 45, 5, 0, 0, 0, 0, 0, 0, 0, 25, 5, 5, 45, 5, 25, 0, 5, 15, 15, 5, 0, 15, 5, 35, 5. The median of the group is 5 and the average is 10. 4. Eight guesses per individual would increase the median by 10 and several individual scores by 30.

Of course, guessing by the human mind is not so mechanical a process as drawing out buttons, but the more complex activity is likely to be the more irregular. It is true that when the individual guesses "No" several times in succession it prepares him for a "Yes"; but this tendency to alternate on the part of the mind is likely to be offset by the fact that when several white buttons are drawn out in succession the black buttons become more numerous than the white. This is not true of the "Yes" and "No" supply. Then, too, mind-sets are likely to increase the irregularity. They are permanent and operate in the same direction throughout the entire test. In all probability the law of chance would not operate again as it did above, but no one can doubt that it would refuse to perform as the advocates of the test in question assume it to operate.

But if one argues that the guessing factor and no more is eliminated by subtracting the wrong answers from the right, one must still further assume that *every wrong answer is a guess;* otherwise, if the wrong answers are subtracted, they will cancel not guesses but actual achievements. That only guesses cause wrong answers no one would care to assume. To justify the test one must assume, in the fourth place, that *just because a certain number of answers are wrong, the same number of right answers must be guesses.* If 10 answers are wrong and 10 right, is one justified in the assumption that the individual tested is in possession of no facts?

But the guessing factor is not the only reason why such a test cannot measure accurately mental ability or achievement. In the first place, the test with its method of scoring credits negative achievement, a kind of achievement that has little, if any, social value. For instance, "No" as an answer to the question, "Is warm air heavier than cold air?" is scored as correct. It shows on the face of it that the individual knows warm air is not heavier than cold air. But does the answer show that he knows cold air is heavier than warm air? His inference is correct if he thinks there is no difference between the two in weight. This objection holds with reference to every question or statement to which the right answer is "No" or "False" if it is used for the purpose of testing information. Is it possible to find out what a thing is by asking what it is not?

In the second place, the test credits dependent achievement. It credits identification of correct answers but not their production. For instance, one might be unable to answer the question, "What are intermittent sounds?" but might be able to identify the answer immediately when the question is stated, "Are intermittent sounds discontinuous?" This objection holds with reference to every question or statement to which the right answer is "Yes" or "True" if it is used for the purpose of testing information.

What then does the final score of such tests represent? No one knows. That it cannot even approximately represent real ability or actual achievement has been shown. Nothing important, therefore, should be done on the basis of the score. Such a test should not be used for determining the mental ages of children, for vocational purposes, for classifying children for school work, for studying individual differences, or for diagnostic and remedial work. The statement that the test should not be used for diagnostic purposes may need further explanation. In a general way such a test may indicate weakness or strength on the part of pupils in a certain unit of work, but this is all it can do. For instance, if 15 answers to a test are right and 9 are wrong, the wrong answers do, of course, show weakness on specific points. But it is assumed that 9 of the correct answers are guesses. Which 9? The teacher cannot determine by the test *what children do know,* and therefore it is not a good diagnostic test.

When used in school as an achievement test, what is its influence on teaching? What influence it has is *reactionary*. In the first place, it cannot test the most important thing about a fact or principle. It is important to know that warm air is not heavier than cold air, but it is much more important that the pupil should know *why it is not*. It is important to know that seaports have a more uniform climate than inland cities, but it is more important to know *why*. There are no "why's" and "wherefore's" in such tests, and without these the school sinks again to the dead level of memorizing facts. In the second place, it does not encourage organization of subject matter. The exercises are not large enough in scope. Most of them call for answers consisting of single facts. A third reason why its influence is reactionary is found in the fact that it does not encourage problem or project teaching. How many of the exercises in such tests would make a good problem or project in the form in which they occur? Lastly, it uses exclusively the form of question that is condemned in nearly every book on pedagogy, the form of question that can be answered by "Yes" or "No" or their equivalents. How would a teacher be rated these days who would conduct a recitation on the pronoun by using such questions as "Does a pronoun express action?", "Is a pronoun a noun?", "Are there four classes of pronouns?", "Are some pronouns called personal pronouns?" There are other reasons, but enough have been given to show that what influence such tests have on teaching is *reactionary*.

And why such tests? The principal reason for their use is to avoid trouble in answering questions and scoring answers. There is no 'royal road' to answering real questions and scoring real answers.

A STUDY IN CORRELATION OF ALLIED ENGLISH ABILITIES

GRACE S. JAMISON
Longwood Commerce High School, Cleveland

Although the following study treats only a very limited portion of an entire field of research it seems important that it be prefaced by a statement of the general significance of all work of its kind. The field of educational research in which these experiments were made is the study of the correlation of mental abilities in various lines of work. This is a field which gives promise of extraordinary reaction upon the efficiency of our teaching in the near future. If the theory of correlated abilities proves true, namely, that a certain degree of capacity for one thought subject (as distinct from mechanical) is proof of capacity for any o ther thought subject, then the old idea of specialized ability for English, for geometry, or for foreign languages is of course exploded. Then the obvious difference of attainment in these various lines by any one student will have to be explained not primarily by special ability for this and inability for that, but by differences in the early preparation of the student for this or that subject, by someone's failure to arouse his interest equally in all branches, or by the unequal success of his teachers in two different subjects at any given time. Now since the efficiency of our teaching, unlike our mental endowment, is something that can be controlled, the time may come when every child's accomplishment in every subject can be raised to its maximum possibility.

It is because the writer is so strongly impressed with the importance of investigating and using the correlation theory, if it proves to be true, that the following study has been made. It is presented here simply as a bit of evidence in addition to that which has been afforded by previous investigations. This experiment measured attainment in two rather closely allied branches, language ability and written composition. It was made by using the Trabue Completion tests for the measurement of language ability, and the Nassau County Supplement to the Hillegas Scale for grading compositions. The crude data used consist of results of tests given in the public school of Republic, Michigan, between September, 1918, and June, 1920, and were collected in the main by Superintendent William L. Connor, now principal of the Longwood Commerce High School of Cleveland, Ohio. During the year 1919-1920 the writer was teaching English at Republic and assisted in giving and grading the tests under Mr. Connor's direction. The writer is indebted to Mr. Connor for the preparation of the graph illustrating this article, as well as for the use of the data.

These data, it should be explained, were not originally collected for the purpose of making this study in correlation. Tests were given from time to time, the records preserved, and graphs of the results posted, with the practical aim of stimulating the work of pupils and teachers. This study in the correlation of written composition and language abilities was an afterthought, and hence the data used have not in all cases been as satisfactory as they might have been if they had been definitely collected for the writer's purposes.

Before going into the results of this investigation it will be necessary to explain briefly the methods of giving and grading the two tests.

METHOD OF TESTING FOR LANGUAGE ABILITY

The Trabue test of language ability consists of sets of incomplete sentences, the required work being to fill correctly as many blanks as possible in a given time. The papers are distributed to the pupils face downward; all are turned up at a given signal so that the test may be accurately timed; at the instant the time is up the command "Stop," is given; and the papers are collected. To insure accuracy in scoring the papers, prepared keys accompany each test. A paper is scored by counting two for each sentence correctly completed; for example, if a pupil had filled the blanks in seven of the sentences and one sentence as completed was unsatisfactory, his score would be twelve. The same test is, of course, never given twice to the same set of pupils. From a little study of such a test it will be seen that the definite amount of language a pupil has learned between the times of two testings cannot be directly measured. What is measured is the degree of intelligence, the quickness of perception and reasoning, and the familiarity with the use of words, which the pupil had at the time of testing. This, of course, shows indirectly his progress in knowledge of language from the time of one test to the time of the next. The great value of the Trabue tests in the present investigation lies in their fairly high correlation with standard intelligence tests. As the results of Trabue tests usually correlate 0.60 to 0.70 (Spearman or Pearson) with intelligence tests, we may conclude that when a set of composition scores correlates 0.75 with a set of Trabue scores the same composition scores would probably have a high positive correlation with standard intelligence tests.

USE OF HILLEGAS COMPOSITION TEST

The Nassau Supplement to the Hillegas scale which was used for rating compositions in this study, consists of a set of ten sample compositions which have been assigned in order the following scores: 0, 1.1, 1.9, 2.8, 3.8, 5.0, 6.0, 7.2, 8.0, and 9.0. On the day set for the test every child in school from the third grade to the twelfth was assigned the same subject, as for instance, "How I Would Like to Spend the First Day of Vacation," "An Exciting Experience," etc. A period of twenty minutes was allowed for the writing, at the end of which the papers were collected and prepared for the scoring which was done by four judges specially selected and carefully trained by study and practice to accurate use of the scale. The names of the writers were removed, leaving only key-numbers to identify the papers. This precaution prevented the judges from being prejudiced by the knowledge that any piece of work was done in any particular grade, or by any certain individual. The papers from all the ten grades tested were next put together and shuffled, and the scoring began. Each of the four judges read each paper independently and determined its score by comparing it with the samples on the scale and assigning it the value of the one it most nearly resembled in merit. Without consulting the other judges she gave it this rating on her score-card, identifying it by key-number only. After the papers had thus been read and rated by the four judges the four scores for each composition were recorded. If three of the scores agreed while the fourth differed from them by not more than two steps, the final score given the paper was that of the three agreeing judges; for instance, if the four values given a paper were 3.8, 3.8, 3.8, 1.9, the final grade was 3.8. If only two of the judges agreed while the third gave a score one step higher and the fourth one step lower than the average score, the average score was the final rating; for instance, if the four scores were 1.9, 2.8, 2.8, 3.8, the final score was 2.8. But if the four scores varied more than those in either of the above instances, the paper was slipped back into the pile to be read and rated again. In passing it is interesting to note that in over 75 percent of the cases the judges agreed, either perfectly or three to one. When all the papers had been satisfactorily rated average scores for grades were determined by reference to the key-numbers, and graphs were prepared to show the progress of each grade since the last testing.

The value of this test as a direct measurement of achievement in the work tested is probably greater than that of the Trabue tests, since any progress in power to arrange material effectively, to present it concretely, and to employ correct technique of expression will be shown directly in the test. The fact that progress in composition ability did actually result during the use of this test is shown by the progress graph (Figure 1). As for the accuracy of the judging, although it is necessarily less rigid than the rating of the Trabue test, at least the method used seems more impersonal and scientific than any other system yet devised for rating compositions.

Thus the two sets of tests were given from which this correlation study is made. Trabue tests were made at the following dates, September, 1918; May, 1919; September, 1919; June,

1920. Hillegas tests were made at the following dates, October, 1918; February, 1919; June, 1919; January, 1920; June, 1920.

Seven Steps in Our Experiment

This correlation study included seven parts which we will first enumerate and afterward describe:

1. The amount of progress in Hillegas scores made in every grade from one date to the next was found and graphed, and the average rate of this progress made by all grades between October, 1918, and June, 1920, was summarized and expressed in percent.
2. The average rate of progress made by all grades in scores on Trabue tests between September, 1918, and June, 1920, was summarized and expressed in percent.
3. These rates of progress in scores on Trabue and on Hillegas tests were compared.
4. The progress of one fixed group of individuals in Trabue and in Hillegas tests from date to date was determined.
5. The rate of correlation was found in each grade between each Hillegas test and every other Hillegas test. (The test given in October, 1918, had to be omitted as individual scores were not kept.)
6. The rate of correlation was found in each grade between each Trabue test and every other Trabue test.
7. The rate of correlation was found in each grade between every Trabue test and every Hillegas test.

Composition progress graph.—The results of the seven studies just enumerated follow in order:

1. The accompanying graph, Figure 1, made by use of the Hillegas score for each grade at each date presents clearly the progress made in proficiency in composition. Besides being graphed this progress was summarized by computing the gain made by each grade between October, 1918, and June, 1920. The average of all these rates was found to be 38 percent, which represents the rate of progress for the whole school between October, 1918, and June, 1920.

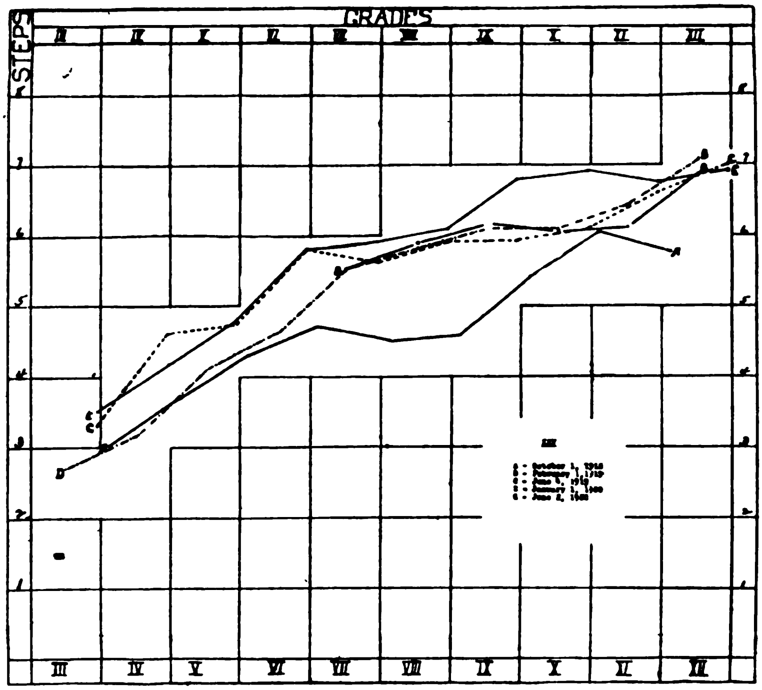


FIGURE 1. COMPOSITION PROGRESS GRAPH

2. In the same manner the average rate of progress in the Trabue scores for all grades between September, 1918, and June, 1920, was found to be 20 percent.

3. Comparison of the progress made in these two abilities during the same period of time may be simply made by inspection of the two percents of gain, 20 and 38. However, the apparent rate of progress in the Hillegas tests with the alternating gains and losses shown by the curves is perhaps more or less invalidated by the fact that a system of special promotions was in use at Republic during the years of these tests. By this system any students showing special ability or making unusual scores in standard intelligence or Trabue tests were permitted to skip a grade while the rest of the class advanced only by normal yearly promotions. This system, it will readily be seen, might frequently lower the score of a given grade at the time of the first

test after a special promotion by removing from it the group of above-normal pupils who had been responsible for the high score of that grade at the last testing. This group of specially promoted pupils being tested now with the grade in which their mental age makes them belong does not materially raise the average score of that grade, while their departure has actually lowered the score of the grade from which they came. Thus an apparent lack of progress shown by graph or table at any time may really be due not to loss of ground but to the correction by a special promotion of a former misleading score caused by testing a group of pupils in a grade too elementary for their ability.

4. *Progress table for one grade.*—Table I, presenting the progress of a fixed group of individuals from September, 1918 to June, 1920, probably gives a more accurate idea of what was happening to language and composition abilities in Republic than the foregoing statements. These individuals were selected from the fifth grade in September, 1918, only those being omitted who happened to be absent from a majority of the tests. This table

TABLE I. SHOWING THE PROGRESS OF TWENTY-ONE PUPILS FROM SEPTEMBER, 1918 TO JUNE, 1920

PUPIL'S NUMBER	TRABUE COMPLETION TEST				HILLEGAS SCALE			
	Sept. 1918	May 1919	Sept. 1919	June 1920	Oct. 1918	June 1919	Jan. 1920	June 1920
1	7	13	9	11	Individual Scores Not Kept	50	38	38
2	4	5	4	5			28	50
3	7	13	14	14		60	50	60
4	8	9	8	9			50	50
5	8	8	12	11		38	38	38
6	4	12	7			38	38	50
7	8	10	8	10			38	50
8	12	12	9	13		50	60	50
9	10	11	8	9		28	38	50
10	10	12	14	11			50	60
11	6	12	12	6		38	38	50
12	7	13	14	10		50	38	50
13	6	14	9	12		50	28	50
14	7	10	11	9		19	38	38
15	8	6	6	6		38	38	38
16	7	10	8	7		28	28	28
17	4	8	12	9		28	19	38
18	2	9	6	10		19	11	19
19	6	12	7	11		50	38	60
20	4	10	9	13		19	11	38
21	4	11	9	10		28	38	38

TABLE II. CORRELATION BETWEEN THE RESULTS OF THE TRABUE TESTS AT ALL DATES[a]

TRABUE COMPLETION TEST	GRADES[b]	TRABUE COMPLETION TEST		
		Sept., '19	May, '19	June, '20[d]
SEPTEMBER, 1918	II	.22	**[c]	**[c]
Test B, Grades II-VII	III	.51	**[c]	.23
Test L, Grades VIII-XII	IV	.33	.53	.48
	V	.43	.77	.23
	VI	.31	.63	.79
	VII	.15	.33	.28
	VIII	.60	.66	.57
	IX	.30	.42	.32
	X	.48	.37	.25
	XI	.25	.40	**[c]
	XII	**[c]	.77	**[c]
	Average	.36	.54	.44
SEPTEMBER, 1919	II		.44	.52
Test D, Grades II-VII	III		.55	.24
Test J, Grades VIII-XII	IV		.32	.38
	V		.55	.33
	VI		.41	.48
	VII		.45	.33
	VIII		.51	.42
	IX		.42	.03
	X		**[c]	.25
	XI		.15	**[c]
	XII		.62	.18
	Average		.44	.32
MAY, 1919	II			**[c]
Test C, Grades II-VII	III			.19
Test M, Grades VIII-XII	IV			.64
	V			.83
	VI			.40
	VII			**[c]
	VIII			.53
	IX			.24
	X			.03
	XI			**[c]
	XII			**[c]
	Average			.41

[a] Table reads, "the coefficients of correlation between the test given in September, 1918, and the test given September, 1919, were: for the second grade, 0.22; for the third grade, 0.51; etc."

[b] "Grade II" means the same group of pupils throughout the time covered by the tests. Other grade names have the same meanings.

[c] Data missing or insufficient.

[d] June, 1920, Test E, grades II-VII; Test K, grades VIII-XII.

shows an average progress in Trabue scores of 48 percent from September, 1918 to June, 1920, and in Hillegas scores, of 46 percent.

5. *The Correlation study itself.*—The correspondence between the work done on Trabue tests at different times was ascertained by calculating the correlation coefficient for each Trabue test with every other Trabue test. The results are tabulated in Table II.

TABLE III. CORRELATION BETWEEN THE RESULTS OF THE HILLEGAS TESTS AT ALL DATES[a]

NASSAU CO. SUPPLEMENT TO THE HILLEGAS SCALE	GRADES[b]	NASSAU CO. SUPPLEMENT TO THE HILLEGAS SCALE		
		June, '19	Jan. '20	June, '20
FEBRUARY, 1919 (Given in High School only)	VII	.89	.86	.91
	VIII	.59	.75	.26
	IX	.18	.50	.48
	X	.80	.60	.14
	XI	.15	.45	.57
	XII	.86	**[c]	**[c]
	Average	.58	.53	.47
JUNE, 1919 Grades III-XII	III		.59	.53
	IV		.87	.87
	V		.53	.85
	VI		.46	.55
	VII		.66	.72
	VIII		.70	.79
	IX		.63	.77
	X		.60	.70
	XI		.49	.66
	XII		**[c]	**[c]
	Average		.61	.72
JANUARY, 1920 Grades III-XII	III			**[c]
	IV			.58
	V			.87
	VI			.66
	VII			.68
	VIII			.83
	IX			.64
	X			.63
	XI			.60
	XII			**[c]
	Average			.69

[a] Table reads, "the coefficients of correlation between the data of February, 1919, and those of June, 1919, were: for the seventh grade, 0.89; for the eighth grade, 0.59; etc."

[b] "Grade II" means the same group of pupils throughout the time covered by the tests. Other grade names have the same meanings.

[c] Data missing or insufficient.

TABLE IV. CORRELATIONS BETWEEN THE RESULTS OF EACH TRABUE AND EACH HILLEGAS SCALE[a]

Trabue Completion Test	Grades[b]	Nassau Co. Supplement to the Hillegas Scale			
		Feb. '19	June, '19	Jan. '20	June, '20
September, 1918	III		**[c]	**[c]	**[c]
Test B, Grades II-VII	IV		.43	.07	**[c]
Test L, Grades VIII-XII	V		.68	.66	.44
	VI		.61	.47	.86
	VII		.27	.41	**[c]
	VIII	.42	.67	**[c]	.41
	IX	.41	.34	.14	.65
	X	.41	.29	.80	.33
	XI	.40	.38	.43	.09
	XII	.39	**[c]	.06[d]	**[c]
	Average	.41	.46	.38	.48
September, 1919	III		**[c]	**[c]	**[c]
Test D, Grades II-VII	IV		.54	.44	.68
Test J, Grades VIII-XII	V		.57	.41	.70
	VI		.75	.50	.48
	VII		.57	.50	.35
	VIII	.64	.76	.69	.39
	IX	.43	.51	.48	.93
	X	.33	.73	.52	.34
	XI	.70	.54	.59	.50
	XII	.12	.19	.34	.28
	Average	.44	.57	.50	.52
May, 1919	III		.47	**[c]	**[c]
Test C, Grades II-VII	IV		.82	.52	.78
Test M, Grades VIII-XII	V		.50	.74	.77
	VI		.48	.12	.56
	VII	.61	.85	.71	.49
	VIII	.31	.36	.77	.43
	IX	.47	.58	.36	.64
	X	.80	.50	.85	.62
	XI	.26	**[c]	.40	.75
	XII	.76	.98	.41	.35
	Average	.54	.61	.54	.60
June, 1920	III		**[c]	**[c]	**[c]
Test E, Grades II-VII	IV		.22	.31	.38
Test K, Grades VIII-XII	V		.67	.57	.65
	VI		.80	.63	.35
	VII		.50	.19	.37
	VIII	.38	.36	.41	.47
	IX	.55	.67	.35	**[c]
	X	**[c]	.56	.65	.64
	XI	.55	.65	.39	.74
	XII	.57	**[c]	.35	**[c]
	Average	.51	.55	.43	.51

---

[a] Table reads, "the coefficients of correlation between the Trabue Completion test as given September, 1918, and the compositions of June, 1919, as measured by the Hillegas Scale were: for the fourth grade, 0.43; for the fifth grade, 0.68; etc."

[b] "Grade II" means the same group of pupils throughout the time covered by the tests. Other grade names have the same meanings.

[c] Data missing or insufficient.

[d] Small class, not taking English in 1919-20.

In determining correlation Spearman's foot-rule formula was used: $1-\frac{6\Sigma D^2}{n(n^2-1)}$, in which D is the difference in rank of the two scores of each individual and n is the number of cases being considered.

6. Correlation coefficients between results of Hillegas tests at all dates were next computed and are tabulated in Table III. In determining correlations between two sets of Trabue tests every score in the class was used each time, but in working with the Hillegas tests the small range of the scores given to the papers in any grade made it more practical to use the method of sampling. That is, three or four cases having each different score were picked (otherwise at random) from the class, making perhaps not more than twelve or fifteen cases altogether, and from these the correlation rate for the entire class was computed.

7. The correlation rates of the Trabue with the Hillegas tests were finally found and tabulated. Table IV shows these results.

CONCLUSIONS FROM DATA OF FOREGOING TABLES

Examining first the correlations of Trabue tests with each other, and of Hillegas tests with each other at two different times, we see that the correlations of Trabue with Trabue average 0.42 and those of Hillegas with Hillegas 0.60. The variation between the respective ranks of the members of a class at two given times is due, of course, to a number of factors, the teaching that has been done between the two tests, the peculiar conditions of health and temper of the individual pupils when the test was given, slight unavoidable differences in the difficulty of two tests intended for the same grade, etc. Thus if the correlation between scores in two Trabue tests or two Hillegas tests is so far from perfect the rate of correlation between a Trabue and a Hillegas test need not be above this rate to indicate the existence of a correlation of ability in language and composition.

In the summary of the correlations between Trabue and Hillegas scores we see that the yearly averages of coefficients in all grades range from 0.38 to 0.61, the lowest being a correlation between two distant dates—September, 1918 and January, 1920—while the highest is of great significance because it is a correlation between two close dates, May and June, 1919. The proximity of the dates May and February, 1919 and of September and June, 1919 gives much significance also to the two high coefficients, 0.54 and 0.57. The average of all coefficients of correlation between Trabue and Hillegas tests is 0.50. If we compare this with the average coefficient 0.42 in the Trabue tests correlated at various dates, and 0.60 in the Hillegas tests we see that there is, according to this study, about the same degree of relation between the work done by the same pupils in language and composition tests as there is between the work done in either of these branches at different times.

In other words our investigation has demonstrated an apparent correlation between the abilities required for these two subjects, language and written composition. Because of the high correlation between the Trabue language test and the standard intelligence tests, pointed out earlier in this article, the significance of the correlation indicated by our study is increased. That is, it would seem probable that if a pupil does somewhat similar work in written composition and in language study, his composition ability corresponds somewhat with his average intelligence, with which in turn his work in arithmetic, history, geography, etc., may correspond. That is, although we have not yet proved the theory of correlation of all abstract subjects our study has at least rendered it more highly probable. We present this investigation, then, simply as one more step in the group of experiments now being made in this important field.

A Few Interesting By-Products of This Study

Meantime the writer wishes to call attention to some of the other results of the investigation, results which have been of quite as much value as the demonstration of the correlation. For instance, as already suggested, the interest aroused in pupils and teachers while these tests were being carried on and the data collected improved the work of the school to an appreciable extent.

This improvement in composition is shown in the progress graph already referred to (Figure 1). Another illustration of the pedagogical value of the correlation theory is the fact, shown by our data, that when the correlation rate is increasing progress is being made in composition scores. That is, while the progress shown by the graph was being made between October, 1918 and June, 1920, the average correlation rate increased from 0.41 for Trabue, September, 1918, and Hillegas, February, 1919, to 0.51 for Trabue, June, 1920, and Hillegas, June, 1920. In other words the conscious efforts of the teachers and the pupils to increase composition ability resulted in higher correlation coefficients.

Again, the value of special promotions was more clearly seen, while these tests were being made, because, as was suggested above, the scores showed the fact that many pupils were in grades where they did not belong. This system of special promotions alone will increase tremendously the efficiency of our public schools in their service both to the abnormally forward and to the abnormally backward children.

We might add to the list of the valuable by-products of scientific testing but we shall conclude with mentioning the effect of a few partial studies of correlation made in Republic from time to time upon the work of teachers and classes, as their desire was roused to show up equally well in all subjects. Apparently the very notion of the correlation theory at once began to increase the amount of correlation already existing. These partial studies gave in Republic a slight indication of the greatly increased efficiency which should result from complete proof and thorough comprehension of the theory of correlation, if these are to come in the future.

Editorials

THE PRINCIPAL AS EXECUTIVE

The finest business men with whom we have come in contact are those who seem to be making the least effort. Whatever may go on behind closed doors, the outward appearance is that of easy mastery of the situation. The art of getting big business done seems to be the art of planning, of delegating, of holding people responsible, and of recognizing ability wherever it is exhibited. To all appearances the executive who has mastered this technique is an interested spectator of a going concern. We suspect that this spectator attitude is more apparent than real, but it is undeniably true that the successful man in large affairs clutters neither his desk nor his brain with unimportant details.

The school man seldom measures up to the standard of the business executive; yet his work involves so extensive an oversight of details and the direction of the activities of so many people that it is in reality comparable to the job of the business executive. Consider, for example, the supervising principalship of an elementary school having twelve to fifty teachers. In this situation the regular teaching corps exceeds the staff of many modern offices. To it is to be added special teachers, the janitor and his helpers, the school nurse, and the school physician. Each teacher has a different job; or in large schools each of several groups of teachers has a different job. For the duties of a first-grade teacher are decidedly unlike those of a fifth-grade teacher or an eighth-grade teacher.

The modern school hums with activity. Bells jangle, messengers scurry to and fro, the tramp of classes is heard upon hall and stairway, the whoop of boys and the scream of girls announce the activities of the playground. The school doctor drops in importantly, not yet quite understanding how it comes about that he is a part of this organization. The nurse requisitions a group of children, the janitor reports, the urbane bookman pays a social call, the superintendent descends, parents, irate and otherwise, storm or argue or plead according to their temperament. The principal's office is the center of it all, and on him rests the responsibility for the smooth running of this machine.

How is he measuring up to the requirements? Is he distracted, wild of eye, abrupt of manner? Does he flit from one item to another, fussily conscientious, undiscriminatingly devoted to each detail? If so, he has something important—important not only for his school but for his own welfare—to learn from the successful business executive.

The fundamental ideas are careful planning, delegation of authority, and firm decision.

We have lately been at no small pains to find out what the current practice of principals is. We have a distinct impression that many of them merely do the next thing. Some apparently realize the value of planning and make fitful attempts to organize their work; but the clamor of immediate events too often beguiles them from the pursuance of their plan with the result that duties as performed bear little resemblance to duties as planned. If you ask a principal what his most important duty is, he will answer "supervision." If you ask him what duty occupies the greatest amount of his time and attention, he will tell you "office work," "correspondence," "overseeing playground activities," "looking after milk for underfed children," "making reports," or "checking attendance."

The superintendent corroborates the principal. Rather extensive reports indicate, as one says, the contrast between the activities of the principal as they should be and the activities as they actually are. Indeed, one superintendent, after listing the five duties of his principals in the order of time devoted to them, disposes of a listing of duties in the order of importance by saying, "Reverse the above list." In other words, in the judgment of this superintendent, the duties to which his principals were devoting the most time were the least important, and those to which they were devoting the least time were the most important.

Evidently, then, a planning of the principal's work with reference to the purpose of the school and of his position is needed to square performance with purpose and to enable the principal to measure up to the size of his job.

It is clear, however, that one may plan skillfully without putting the plan into effect. As we see it, the greatest obstacle here is the failure on the part of the principal to utilize the abilities of his colleagues. Principals report an enormous amount of time devoted to inspecting buildings, supervising playgrounds, checking up attendance, policing halls, issuing library books, running athletic events, carrying out special campaigns or "drives," organizing entertainments, and handling textbooks and supplies. These are not principals' duties if they can possibly be avoided. Indeed, some of them are of such a character that they should not even be attempted if they cannot be delegated. In our conferences with principals and in our examination of reports from them we are impressed with nothing more distinctly than with the fact that principals are refraining from utilizing the human energies at their disposal and that they are thus depriving their schools of the educational leadership which the principal's position essentially contemplates.

We have mentioned firm decision. In school matters as in business matters it is an understood principle that it is often not so important how a thing is decided as it is that it be *decided*. Any enterprise will suffer if decision is too long pending. Misdirected energy, wabbly purposes, and a general lack of "tone" surrounds the executive who habitually hesitates and postpones the day when he must make his decision.

And when the decision is made, there must be no doubt as to its nature or as to whether it is in effect. Principals often announce that certain things must be done; and then if there is opposition, or even if there is not, they allow the procedure to be modified or abandoned with the lapse of time. No principal can afford to permit this laxity. Either a rule of action is in force or it is not; and as long as it is in force, it should not be evaded. If, after the announcement of the principal's decision, it becomes evident that the decision should be modified or withdrawn, it is the principal who should make the change, and he should be man enough not to hesitate to reverse himself. His authority will be far better respected if he makes the change himself than it will be if the change is made by allowing teachers and pupils to disregard instructions.

In the schools we have, therefore, an abundant opportunity—indeed a crying need—for the application of business practice. The principal, to say nothing of the superintendent, is a real business executive. It is true that his business is peculiar, but it requires those elements of planning, of delegation of authority, and of decisive action which characterize the best type of industrial and commercial leaders. In some way the principal must obtain a clearer vision of his job. Moreover, the conditions at headquarters must be made right for him to exercise the functions of his job. As things are now, relatively few schools are obtaining from their principals either the amount or quality of educational leadership to which they are entitled.

B. R. B.

PROGRESS AND RESEARCH

If the modern movement in educational research is satisfying our doubts in many directions, it is equally true that it is raising new doubts in other directions. In order that there might be an institution such as the school, and in order that it might be developed to its present proportions, it was necessary that decisions should be made and procedures adopted without waiting for evidence.

Now, however, that richer and more diversified means of evidence are at hand, it is natural that some of the old suppositions and some of the accepted procedures should be subjected to scrutiny. It has been widely assumed, for instance, that kindergarten instruction has sufficient value to justify large expenditures for its maintenance. Yet, we really do not *know* whether this money is well spent or not. Each of us is limited by his own small list of cases and his insufficient examination of them. Again, superintendents have adopted departmental teaching (and abandoned it, too, for that matter) on the basis of some one's say-so. There is no competent evidence to prove whether a departmental organization is better than a grade organization or whether the reverse is true. Does anybody know whether vertical or horizontal supervision is best, or are these installed merely on the basis of a "hunch," in other words (to adopt the lingo of the craft), on the basis of *a priori* reasoning?

Nor do we need to look into the past. Our up-to-the-minute literature is besprinkled with reports of the grouping of children on the basis of intelligence. Yet no one has proved that classes consisting of pupils of approximately the same mentality may be more effectively taught than classes in which bright, average, and dull pupils are represented. The case for homogeneous grouping looks good, that is to say, it is plausible; but in a scientific sense we know nothing about it. Nobody seems to have thought to compare even the measureable achievements of pupils of equal mentality in homogeneous and heterogeneous groupings.

Supervised study and the project method, unilateral lighting and teachers' retirement systems, serial bonds and school-board independence should be re-examined with reference to the results they secure. An imposing array of research problems is involved in this re-examination.

Such research is of evident value. Those who confine their attention to tests and more tests are narrowing their sphere of usefulness in education. Tests are means to larger ends; and although there is a place for work directed toward the perfection of tests, there is no place for "just giving" tests. Progress depends first upon clear thinking as to the purposes of the school, and second upon scientific inquiry into the worth of the means proposed or already adopted for accomplishing these purposes.

B. R. B.

Reviews and Abstracts

FRASIER, GEORGE W. *The control of city school finances.* Milwaukee: The Bruce Publishing Company, 1922. 132 pp.

Two criteria are used for classifying cities in regard to their school finances—namely, the method by which school board members are chosen and the degree of financial independence regularly possessed by the board of education. The author has data for 169 cities which he has divided into nine classes. These classes differ from each other in the two respects just mentioned. For example, the first class consists of cities whose board members are elected by the people and in which the board has financial independence—i. e., is able to levy its own taxes and expend the proceeds of them. At the other end of the line, the ninth class is composed of cities in the benighted condition of having a board appointed by the mayor which is entirely dependent upon the municipal government for its finances. The classes in between the first and ninth have varying degrees of independence, some with elected and others with appointed school boards. These classes are arranged in the order which the author conceives to be theoretically most favorable for school support, class one being the most favorably and class nine the least favorably situated.

The author has devised an index number by means of which he seeks to appraise the school systems in these 169 cities. His next step is to find out whether the efficiency of cities tends to increase or decrease from cities of class one to those of class nine. He finds a distinct relationship of this sort, the correlation coefficient being minus 0.27. By this and by other means he concludes that the city with an elected board and with financial independence for the board has a better chance to operate an efficient school system than do cities not having these advantages.

B. R. BUCKINGHAM

Ohio State University

HERRING, JOHN P. *Herring revision of the Binet-Simon tests.* Yonkers: World Book Company, 1922. 56 pp.

The Herring Revision of the Binet-Simon Tests is an individual examination for the purpose of measuring mental ability. Its purpose is to supplement the various other revisions which are becoming more or less familiar to school children. The questions are asked and for the most part answered orally. Many of the exercises are similar to those in the Stanford Revision. The mental age derived by means of the Herring Revision has the same significance and meaning as the mental age by the Stanford Revision.

The examination consists of thirty-eight tests. A score on the test as a whole is the sum of the scores obtained for the separate tests. Provisions are made also for obtaining the mental age by use of fewer than the thirty-eight tests. Tables are given which show the mental-age equivalents of each total score. The manual contains practically all of the supplementary material used in testing the children and also a set of standardized answers for which credit is given. For example, the picture and color cards are included in the manual itself. Also the articles which the pupil is expected to name are given in picture form in the manual so that the examiner does not have to carry around such things as a hair pin, piece of string, a nail, etc. The complete record of each pupil may be registered on a single card and may be filed for further reference.

Dr. Herring has succeeded in making up a manual which may be readily used by the experienced tester. If the norms which he has derived prove to be reliable, he has made a valuable contribution to the testing of individual children. Many of the items in the Stanford Revision have been published in newspapers and popular magazines so that many children have become familiar with them. It is also true that children have been tested several times with the same form and consequently have obtained practice on the given material. The writer believes that there should be two or more forms of individual tests of intelligence just as there are two or more forms of the better group tests of intelligence.

P. R. STEVENSON

Ohio State University

DEARBORN, W. F., INGLIS, ALEXANDER, and others. *Psychological and educational tests in the public schools of Winchester, Virginia.* Charlottesville, Va.: University of Virginia, 1922. 54 pp.

This investigation resulted from the survey of the Winchester public schools which was made by the General Education Board of New York in 1918. Its purpose, as stated in the Foreword, was "to institute an intensive study of not only the methods of education in the Winchester schools but to pay especial attention to the achievements and capacities of the school children." Dr. Dearborn and Dr. Inglis of Harvard University were asked to plan the investigation; the University of Virginia, the Virginia State Department of Education, and the staff of the public schools of Winchester cooperated in carrying out the work.

The methods adopted are set forth in Chapter II. The general capacity of the pupils was determined by the Dearborn Group Tests of Intelligence (preliminary form), the Army Alpha Test, and, in extreme and doubtful cases, the Stanford Revision of the Binet-Simon Scale. The achievement of the pupils in school subjects was determined by the Courtis Standard Arithmetic Tests, the Monroe Silent Reading Test, the Virginia Reading Test, the Ayres Spelling Test, and the Holmes Handwriting Test. For purposes of checking the teachers were asked to give judgments of each pupil's intelligence, scholarship, and industry.

The discussion of results and conclusions are given in Chapters III and IV. It was found that pupils are permitted to enter school at almost any age, thus causing very heterogeneous groupings and preventing the school from giving as much education as it should. The psychological tests showed that there was very little relation between grade classification and mental ability. The results of the achievement tests showed the same; some children were placed in the upper grades who had not had even the rudiments of the subjects. As to teachers' judgments, two conclusions are drawn: (a) that teachers differ in their ability to estimate capacities and achievements and (b) that, in a large proportion of the cases, teachers' judgments cannot be relied upon.

The general organization of this report is excellent. There are detailed analyses of test results, but they are placed in an Appendix and do not interfere with the discussion, which is simple and direct.

W. W. COXE

Ohio State University

BALDWIN, BIRD T., and STECHER, LORLE I. *Mental growth curve of normal and superior children.* (University of Iowa Studies in Child Welfare, V. 2, No. 1.) Iowa City: State University of Iowa, 1922. 61 pp.

This monograph reports the results of re-tests with the Stanford Revision of the Binet Scale on a group of 143 normal and superior children of the ages 5 to 14. All 143 children were tested twice, 87 were tested three times, 78 four times, and 35 five times. The median intervals between tests were about 15, 13, 7, and 6 months, respectively. The tests were made by four different examiners; the correlations between successive tests ranging from 0.72 to 0.93.

The results show that the I. Q. curve is practically a horizontal line for both superior and normal groups as well as for the sexes taken separately. Apparent exceptions are found in the sharp rise at the ages 11 to 12 for both superior boys and girls, and at 12 to 13 for the normal girls. These increases resemble closely the early adolescent spurts in physical-growth curves. It is difficult to decide the exact statistical significance of these differences because the probable errors of the differences were not stated.

The mean I. Q. was found to increase from test to test, the greatest increase occurring in the third to fifth examinations. This is probably due in considerable part to the fact that the last three tests were made within a period of about 13 months. The average error of estimate of a second I. Q. from an obtained value equalled 4.7 ± 0.5.

The authors conclude that the present Binet Scale is not well suited to the measurement of mental growth. The need for duplicate scales is clearly indicated. The reviewer, however, feels that much of the personal equation evident in the present study (e.g., re-tests by one examiner gave an average gain of about 4 I. Q. points although another showed an average loss of more than 1 I. Q. point) could have been eliminated by the plan of using complete stenographic records scored independently by several psychologists and the results averaged. Systematic errors due to the personal equation of the examiners are not adequately represented by coefficients of correlation.

In general this study makes a most welcome contribution to our knowledge of mental growth. The inclusion of a normal and a superior group is particularly valuable.

G. M. RUCH

Stanford University

WILSON, G. M. *Connersville course of study in elementary mathematics.* Baltimore: Warwick and York, 1922. 110 pp.

The "Connersville Course of Study in Elementary Mathematics" was prepared ten years ago by G. M. Wilson, then superintendent of schools at Connersville, Indiana, in conjunction with his corps of teachers. A revised edition has recently been published by Warwick and York.

It is a small book of few more than a hundred pages, the first score of which are devoted to an "Introduction" and to "General Directions and Suggestions." The "Introduction" is concerned with the purposes of arithmetic, the requirements of a course of study, how to make a course of study, and how the subject matter was selected. These first two sections make the book worth while for they are full of helpful suggestions. The remainder is devoted to a detailed discussion of the work by grades.

On the whole the course of study is carefully graded with special emphasis on oral work, on fundamental processes, and on making these processes automatic. The elimination of lengthy explanations and of unnecessary subjects, the excellent continuity of the work from grade to grade, the uniformity of process, and the review of work of former years, are among the outstanding strong points which impress one after reading the book. On the other hand the noticeable weak points are the lack of suggestions as to how the teacher may give real meaning to abstract processes when first introducing them; the problems given for the lower grades are of no real interest to the children; and not enough emphasis is given to learning combinations, tables, etc. through "doing." Since this book was first written ten years before the report of the National Committee on Mathematical Requirements was prepared, it is not surprising that the first edition was not in accordance with the recommendations of the National Committee. It seems unfortunate, however, that the recent revision should have almost completely disregarded its recommendations. For example, the simple algebraic equation is introduced through an arithmetic problem, whereas the Committee distinctly recommends that observational or intuitive geometry should precede the introduction of algebra and should be the means of its presentation.

The book is well worth while not only for a teacher to read but also to study and use as a handbook. It has an abundance of suggestions, many of which are helpful and profitable. In reading the book one is particularly impressed with the thoroughness of the study and preparation needed by the teachers for their task. What an awakening and professional growth must have come to the entire corps of teachers from this participation in so excellent a project!

MADGE PERRILL

Columbus Public Schools, Columbus, Ohio

DETROIT PUBLIC SCHOOLS. *A Course of Study in Arithmetic,* 1921, 24 pp. *A Course of Study in Geography,* 1921, 160 pp. *A Course of Study in Spelling,* 1921, 80 pp. *A Tentative Course in Music,* 1921, 54 pp.

The above mentioned courses are issued in separate pamphlets. Each covers the printed outline of work throughout the elementary schools. Mechanically they are well executed, being printed in good, clear type on unglazed paper and attractively bound. Careful attention has been given to the composition with the result that the organization of each pamphlet is clearly evident as soon as one begins its use.

ARITHMETIC

The course of study in Arithmetic is of the skeleton type. Little or no help is provided by showing how the types of work indicated may be motivated through discovering social situations in which they are serviceable. Not only is emphasis placed upon thoroughness and accuracy, but the aim is to require only that which will be of use. Throughout the course "Required," "Recommended," and "Optional" work are clearly distinguished. Those examples which are not considered worth while are carefully indicated under the general title of work to be "Omitted." There are internal evidences that three different groups of teachers prepared the outline of work, and that the three groups did not check their point of view.

GEOGRAPHY

The course of study in Geography was prepared by two groups of committees, one responsible for the third, fourth, and lower-fifth grades, the other for the upper-fifth and sixth grades. The first twenty pages are devoted to an introduction, a discussion of aims, and method helps. One finds under the head of "aims" the following, which properly belong in the technique of teaching and learning: "visualize maps," "gather geographical material," "organize geographical material," "acquire general information," "discover problems," "solve problems," "study pictures," and so on. The term "aims" has come to stand for much larger things in education than those which are so labeled in this course. The method helps on "how to develop a topic," "how to solve a problem" are brief but clear. The outline by grades on page 21 makes clear the responsibility of each class.

There is little to suggest relative values. The impression is apt to be left that all the work listed to be done should be covered. Of particular importance, however, are the definite references to sources of help in developing the various problems and topics indicated throughout the course. From the fifth grade on, the work to be done is more generally indicated in terms of problem statements than obtains in the earlier grades. A strong feature of the course is the indication by grades of the minimum essentials in location and information. The closing page stresses the use which should be made of the material which may be borrowed from the Children's Museum.

One misses in this course a statement of the large social objectives which should be achieved through the teaching of geography. While the problem of indicating the material to be covered is an important one, yet one regrets the absence of those helps, directions, and interpretations which should guarantee large socializing and humanizing values to the children through the study of geography.

SPELLING

The course of study in Spelling is worked out with greater thoroughness and with more evidences of a consciousness of the social need of teaching this subject than is manifest in the other monographs. The statement of the spelling problem by S. A. Courtis, in which he emphasizes the importance of developing the "ideal of not making mistakes in spelling," and of placing more emphasis "upon purposing than upon spelling" launches the work throughout on a high level and with good inspirational effect.

The monograph is divided into nine parts. Part I is the teacher's manual for the first four grades, and contains definite, helpful suggestions as to how to go about achieving the purpose of spelling. Part II gives the number of 100 percent words in the first four grades, the number of words required varying from 50 in the I-A grade to 110. Part III gives the supplemental list for the first four grades. Part IV provides the teacher's manual for grades five to eight, and is just as replete with excellent suggestions as is the manual for the first four grades. Part V indicates 100 percent words for these grades. Part VI is the supplemental list for these grades. Part VII contains dictation tests. Part VIII, devoted to the use of games, is an excellent discussion of how to motivate the teaching of spelling. Part IX is a well-selected bibliography. This is the best course of study in spelling which it has been the reviewer's privilege to examine.

MUSIC

The course in music opens with an introductory word to teachers by Superintendent Cody, followed by a two-page discussion of method by S. A. Courtis, in which he shows the large values which may come to children by the study of music through proper and high "purposing" in their study. A committee of twelve persons assisted those who were in charge of the working out of the course. This manual is full of suggestions to teachers, and is calculated to make their work easy. The material suggested for singing is definitely indicated as to sources. It is classified under such headings as occupations, toys, pets, nature, baby, play time, holidays, mother goose, and so on. Precise directions as to the elimination of monotone singing in the lower grades are provided. How the music readers are to be used and the attention to be given to sight reading are carefully indicated. The last eight pages are devoted to a phonograph record index by grades.

These courses of study represent a high order of effort. Particularly do those in Spelling and Music evidence a definite grasp of the social purposes of the school. Each is worked out with sufficient detail for the teacher to know what her responsibilities are; and a large variety of helpful suggestions and directions are right at hand for her use.

H. B. WILSON

Superintendent of Schools, Berkeley, California

Betts, G. H. *The new program of religious education.* New York: The Abingdon Press, 1921. 105 pp.

Facing the title-page of this book is a quotation from President Harding which constitutes a sort of text for the discussion which follows: "Childhood is the basis of the future, and I believe in religious instruction for American children. The future of the nation cannot be trusted to the children unless their education includes their spiritual development. It is time, therefore, that we give our attention to the religious instruction of the children of America, not in the spirit of intolerance, nor to emphasize distinctions or controversy between creeds or beliefs, but to extend religious teaching to all in such form that conscience is developed and duty to one's neighbor and to God is understood and fulfilled."

Professor Betts addresses himself to church folk in this little book, which presents a lucid statement of the principles underlying the present widespread movement for more effective religious education, and constitutes a straightforward and stirring challenge to the churches to conceive their own life and service in educational terms. A sharp and detailed contrast is drawn between "religion through education" and "religion through evangelism"; the church's neglect of religious education is pictured; and the changes in principle and method which are required for the adoption of a genuinely educational program are briefly, but convincingly, set forth.

Such a popular statement of principles has been much needed, and the book will be useful to many. The author somewhat exaggerates the neglect of religion by church colleges, I think. I regret, moreover, his choice of the words "evangelism" and "evangelistic" to denote the mistaken policy of reliance upon reclamatory conversion which he very properly contrasts with the policy of Christian nurture and education. "Revivalism" and "revivalistic" would be better; or perhaps other terms could be found. Evangelism ought to retain its primary meaning, as a bringing of the evangel, the gospel of Jesus Christ, to bear upon human life in redeeming power, by whatever method. Not evangelism or education, but evangelism through education, should be the aim of the Christian church.

L. A. Weigle

Yale University

Boraas, Julius. *Teaching to think.* New York. The Macmillan Company, 1922. 289 pp.

Educators have long been accustomed to hear that our schools do not teach real thinking. Of late there has been a tendency to admit the truth of the accusation and at the same time to shift the stigma from the educative process to the nature of the material with which it works. Only a few pupils, it is asserted, have the native capacity to do their own thinking. Let these few be trained to think for themselves and for everybody else too, and let the "vast residuum" be trained to follow, "believing where they cannot understand." In the hands of injudicious supervisors, mental testers, and biological determinists, the element of truth which the doctrine of individual differences contains seems in danger of being magnified in application until the American school system becomes a Prussianized machine, the quintessential embodiment of that educational determinism against which Bagley recently sounded his timely though rankling warning. The reviewer cannot but feel that Professor Boraas' book is a salutary influence toward mitigating the mechanistic tendencies which have always been the bane of education and which the present test-and-standardization movement, with all its virtues, seems in danger of re-enforcing.

In the face of this danger it is refreshing to find an author who says that "we need a Socrates in every schoolroom" (p. 3), and who contends that the development of ability "to think honestly, persistently, and effectively about the important problems of life" (not the ability to put the right word in the right blank) is "The Greatest Thing in Teaching." The words just quoted form the title of the first chapter in which is the following significant passage: "Perhaps the reason why so few of us do our own thinking is that the schools never seriously attempted to teach us to think for ourselves. Perhaps the reason why so many of us are blind followers is that we have been trained to be blind followers" (p. 4).

The author does not stop with a mere "perhaps." He presents definite evidence of the widespread failure of our schools—not their failure to *accomplish* training in thinking, but their failure *to make any effort to accomplish it.* (See especially pp. 5-12.) We have said that the majority of our pupils are incapable of thinking, but we haven't given them a chance to show whether they are or not.

The purpose of our author is "to discover the principal types of thinking which are required in everyday life and to indicate practical ways and means for their development in the ordinary school" (Preface). Accordingly, he analyzes thinking into such types as "Individual Judgment," "Co-operative Thinking," "Imaginative Thinking," "Critical Thinking," etc., and devotes at least one chapter to each type, showing its value and giving definite suggestions, some original and some borrowed from acknowledged sources, for its development in the school. Most of these suggestions will impress the teacher as worth trying. A feature common to them all is an emphasis upon the position of the teacher as cooperator, not as subaltern of a despotic hierarchy, a little lower than the supervisor, crowned with data-sheets and scales. A teacher who gets into the spirit of this book will have an inspiring sense of being engaged with his pupils in the great joint enterprise of adding to the worth and meaning of life by the application of sympathetic intelligence to its myriad problems.

This spirit which animates the book, the eminently readable style in which it is written, and the many practical suggestions it contains go far toward atoning for its faults. Feeling, as one does, the generous, altruistic spirit of the author, one can forgive him for such lapses as he exhibits in his suggestions on "How Not to Learn to Do Things" (pp. 27-28). Here and in some other places he slides into a jejune wordiness which rather amusingly suggests the "efficiency literature" with which book-agents are trained. There are other things to forgive. Many readers will with reason object to the looseness of such analyses and classifications as appear on pp. 29, 43-45, 167-168. The author himself seems to be aware of this looseness and not to be greatly concerned about it as his purpose is not to formulate water-tight principles but to radiate suggestiveness. One gets the impression that the author's background of logic is deficient. For example, his realization that induction and deduction are not two fundamentally different modes of thought is excellent, but his attempt to synthesize them into a statement of generalized procedure covering all thinking is noticeably vague and insufficient (pp. 166-168). When, as in Chapter II, he tries to give a close-knit analysis of the thinking process, he makes some excellent statements about its problematic and experimental quality, but through some indiscretion he is led to define it as "an inner behaviour"—a phraseology which carries objectionable Cartesian associations which our author himself would probably be quick to repudiate.

For the background of cogent logical theory in which this book is somewhat lacking the teacher will do well to look to Dewey's "close-written, close-meditated" *How We Think* and to the superb brief treatment of the thinking process which is to be found in Professor Bode's recent *Fundamentals of Education*. These two books should form a part of every teacher's professional library, and to them I do not hesitate to say that *Teaching to Think* should be added. The book seems to me to have the faults I have mentioned, but they are more than compensated by the many applicable suggestions it contains and the fine human spirit by which it is pervaded. That spirit is of the essence of progress, and its diffusion is a part of the project to which every teacher should be dedicated—"the project of making life liveable and joyous."

N. S. MADDOX

Ohio State University

BLACKBURN, MARY. *Montessori experiments in a large infants school.* Introduction by Edmond Holmes. New York: E. P. Dutton & Co., 1921. 143 pp.

This book offers suggestions for the practical application of the Montessori method in large classes of young children. Progressive educators realize the necessity for physical freedom and for work adapted to mental capacity, especially for little children. It is difficult to provide the better type of education in the present prevailing school conditions of crowded classes. Anything that helps towards solving this problem, as Miss Blackburn's book does, is a positive contribution towards current educational literature.

Miss Blackburn has a keen sympathetic understanding of the nature and needs of small children. She desired a method by which the initial steps in school training could be taken in a normal, childlike way. She felt that in acquiring subject matter each child should be permitted to advance at his own rate, that the brighter pupils should not be hampered by being held back for the slower ones and that the slower pupils should not be continually prodded and blamed for not progressing faster than they had ability to advance. The result which she desired even more than progress in school subjects was that the children should be free, happy, busy, self-reliant, helpful.

For four years Miss Blackburn experimented with the Montessori method and materials. She describes economical and easy ways of making Montessori materials from toys and other materials easily procured. She suggests simple objects which can be added to the Montessori materials or substituted for them. Suggestions are also given for storing the apparatus.

During the experimenting in the use of the materials she found it advisable to change the character of the class organization. Miss Blackburn divided her children vertically into groups of a certain number instead of horizontally into those of a certain age. Each class included children between 3 and 7 years of age. She found this more convenient for it was possible to keep all of the children busy with one set of apparatus. There was also a distinct advantage in the opportunity which it afforded for social development; a more normal situation, a spontaneous social life, was provided for the children; they could help and be helped. There was also more opportunity for the teacher to give her attention to individuals at the time when they needed the assistance, at other times the children could develop initiative and independence by working alone. She gives suggestions for keeping the records of each child's progress.

Three chapters in the book are devoted to experiments made in reading and writing and two chapters to number work. These particular chapters contain only a few suggestions for American teachers. Number work here is postponed until the later grades. Children acquire what is necessary through actual experiences or score games.

The book as a whole is an account of a successful attempt to carry the thoughtful, experimental spirit of Mme. Montessori into a new situation. Miss Blackburn shows herself a creative follower of Dr. Montessori and as such wisely fitted to lead others. As Dr. Holmes states in his introductory chapter "the highest function of the Montessori teacher is to give her pupils freedom for self-development, to give it judiciously, sympathetically, tactfully, fruitfully. But if one is to give freedom to children, young or old, one must first claim freedom for oneself."

LUELLA A. PALMER

Director of Kindergartens, New York City

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors. Address all correspondence concerning this department to Doctor E. J. Ashbaugh, Ohio State University, Columbus, Ohio.

Mr. W. H. Hughes, formerly district superintendent at Claremont, California, has been appointed Director of Educational Research and Guidance in Pasadena.

Dr. W. W. Theisen, formerly Director of the Bureau of Research, Cleveland, Ohio, has become Assistant Superintendent of Schools in Milwaukee. His work will be devoted to the secondary field and will not involve the general field of research activities as we might have expected.

Mr. G. W. Frasier, Director of the Department of Classification and Statistics in Denver writes that on the basis of a thousand cases of first-grade children tested with the Detroit, Dearborn, and Cole Intelligence Tests the Cole Test showed a higher correlation with mental age as determined by the Stanford Revision than either of the others. The Detroit Test failed to distribute the upper ranges of intelligence; while the Dearborn distributed the upper ranges better than the lower.

Numerous attempts have been made to devise a form for keeping permanent individual record of all the tests which a pupil takes during his school career. One of the most interesting of such forms that has come to our attention is one lately received from Logansport, Indiana. It was transmitted by Miss Myrtle L. Kaufmann, but Miss Kaufmann was so busy when she sent it that she forgot to write a letter and say whose work it represented. At any rate the card is compact and seems to be adequate for the purpose.

The University of Michigan, Bureau of Educational Reference and Research, reports in its Bulletin, dated September 7, 1922, the results of its testing program conducted last May. The school subjects considered were spelling, reading, and arithmetic; and the same tests were used in May as had been used the previous October.

The results for May indicate that the Michigan schools are not yet quite up to the standards given by the authors of the tests. Achievements, however, are much better than was the case the previous October. "The amount of improvement . . . is at once startling and shows that an unfavorable situation can be improved if it is discovered and if attention is given to it." This, we take it, is the lesson to be derived from the Bureau's May test. The important consideration is not where a school stands but the direction and rate of its improvement.

We have just received from Superintendent J. W. Studebaker of the public schools of Des Moines, Iowa a mimeographed pamphlet which his office has issued to members of all committees involved in textbook adoptions. The pamphlet includes (1) statement of procedure to be followed, including directions and rules for the guidance of committees; (2) sample score sheets in biology, geography, and history; (3) guiding principles as accepted by a course of study committee for geography, and requirements for construction and printing of books; (4) the same type of material for history as is given in (3). The tentative score card which has been developed covers the major items of interest, comprehension, permanent value of subject matter, value of method, and mechanical elements. These are subdivided and the various items weighted so as to give a total score of one thousand points. The score sheets and guiding principles enunciated should be of much value to committees elsewhere who are working on the problem of scientific textbook adoptions.

Mr. E. P. Branson, Director of the Department of Research, Long Beach, California, has sent us a report of Courtis Arithmetic Tests in his Bulletin No. 1, September, 1922. The results of the tests in comparison with May, 1921, show losses in grades v and vi, and rather consistent gains in grades vii and viii. The results also show that there is practically no change in variability in rate from grade to grade, but there is a marked decrease in the variability of accuracy.

One of the especially pertinent statements contained in the Bulletin is as follows: "It can not be asserted that to fail a pupil is to lower the quality of his future achievement, but it may safely be asserted that special coaching, attractive drill devices, exact knowledge of the improvement that is being made, and much motivation are needed to prevent a probable lowering of quality of work after non-promotion."

We are in receipt from the World Book Company of copies of the Thurstone Employment Test for Examination in Typing and Examination in Clerical Work. The test in clerical work involves checking for errors in simple fundamental operations in arithmetic, checking for misspelled words, cross-out letter test, a substitution of digit for letter test, two types of filing tests, a test in simple arithmetic problems, and a proverbs test.

The typing test involves re-typing of an author's corrected copy, a tabulation of some itemized material written in long hand, and a checking test for misspelled words. The World Book Company has also brought out the

Downey Will-Temperament Test in group form. The test is very similar to the author's individual test, but is so arranged as to make it possible to score on the basis of the amount of work done in a prescribed period of time instead of the time needed to perform a prescribed task. Two items of the individual form could not be adapted for the group use, and substitutions have been made. A test for self-confidence has been substituted for a resistance to opposition; and a non-compliance test for a reaction to contradiction.

Readers of the Journal will be glad to know that the publication, *School Life*, is to be resumed. The September number will be issued shortly and it will continue to appear monthly, hereafter, except July and August. From a letter sent out by Commissioner Tigert, we find certain changes in the rules governing its distribution for "under the terms of the Public Resolution authorizing the resumption, those who receive the publication must pay part of the cost; an amount must be charged for subscription equal to the cost of printing, plus 10 percent. This will not cover any part of the cost of editorial work. The publication carries no advertising. The price of subscription has been fixed at 30 cents a year; but 25 copies or more will be sent to the same address at the rate of 25 cents a year for each copy.

"Persons who wish to receive *School Life* regularly should forward the subscription price at once to the Superintendent of Documents, Government Printing Office, Washington, D. C. Please do not send money to the Bureau of Education."

Professor T. H. Schutte, director of the Training Department, State Teachers College, Moorhead, Minnesota, has devised a very complete plan for rating teachers. He does not pretend that it is an absolutely scientific scale, but offers it in the hope that it will (1) aid the teacher in analyzing herself with view to improvement, and (2) aid the training officer in his analysis of the teacher with a view to her improvement.

Rating schemes are bound to be with us. There is a definite tendency to relate schedules to teaching merit. In this connection we have lately received a plan from Chester, Pennsylvania, to the effect that the new Pennsylvania law requires a rating of teachers as a basis for the application of the salary schedules.

Whether or not one is in a situation for rating his salaries directly, one will certainly be interested in the "Moorhead One Hundred Point Scale for Rating Teachers." It may be obtained from Professor Schutte.

The Bureau of Educational Research of Ohio State University gathered material last summer in connection with a course for elementary-school principals which the director was offering. Superintendents, principals, and teachers were asked to list the duties which were actually being performed by supervising principals in their locality, and to present these items in the order of the amount of time devoted to them by the principal in question. They were also asked, as a second part of the question blank, to indicate the duties which, in their judgment, *should* occupy the attention of principals, and to arrange them in the order of their importance. Not many responses were received from teachers, but it was remarkable how frequently both superintendents and principals practically inverted their lists. In other words, the duties on which the most time was actually being spent often proved to be the least important in the judgment of our correspondents. One superintendent, after giving a list of five duties in the order of the time spent on them by his principals, answered the question concerning the important duties by saying "reverse the above list."

If these indications tell a true story, what are we going to do about it? What, in particular, are superintendents and principals and Boards of Education going to do about it? But the responsibility for this condition does not rest wholly—perhaps not mainly—with the principals themselves.

Superintendent C. F. Ridgley of Logan, Ohio, has been doing some careful testing of the pupils of his schools. Most of his teachers have been studying under Professor Morton of Ohio University. Accordingly, the tests were more adequately given and more helpfully analyzed than is usually the case. Superintendent Ridgley has issued two mimeographed reports in which he presents his results. He indicates that the following purposes were in mind throughout the testing:

"(1) To discover the location and extent of pupil misplacement in the Logan schools; (2) to discover the abnormally bright and abnormally dull pupils; (3) to find a more equitable basis upon which teachers may compare results in achievement."

The Illinois General Intelligence Test was given in grades III to IX, and later in the year the Haggerty Delta I was given in grades I to III. Achievement tests in several school subjects were also given.

While the tables in which the results of the tests are displayed are often of the conventional kind, there is a highly suggestive relating of mentality to grade placement and to achievement. Superintendent Ridgley is not sparing in criticism of his own organization. For example: "It is easily seen, then, that our promotions in Logan have been influenced more strongly than we realized by chronological age considerations; that there is more holding back of pupils than acceleration; that the compactness of distribution of retardates indicates rather uniform promotion after a certain time in a grade regardless of accomplishment."

In a recent letter Superintendent Ridgley indicates his intention to carry out this year as extensive a program of testing as he conducted last year. He states that great benefit was derived from the work last year and adds, "We have planned to continue this program so long as we seem to profit by it."

Rochester Chart for Attainment in Reading

We have lately received some interesting material from Rochester. Our readers will recall the article by Assistant Superintendent O'Hern in our March, 1921 issue on "The Development of a Chart for Attainment in Reading." Principal William E. Hawley of the Francis Parker School at Rochester wrote an article which appeared in the April, 1921 issue on "The Effect of Clear Objectives on the Teaching of Reading."

In these articles it was apparent that the people at Rochester, although they had worked assiduously in bringing out their attainment chart, did not by any means regard it as in final form. A part of the material lately received from Rochester consists of an extensive revision of the reading chart. It is our judgment that, good as the first chart was, it has been greatly improved in the present edition.

It is clear that if definite suggestions produce good results in reading, they may be expected to produce good results in other subjects. At least this appears to have been the line of thought of the school people at Rochester. Accordingly they have brought out a chart of attainment in arithmetic and a similar one for minimum requirements in phonics. These charts have been prepared by the Department of Elementary Grades and Kindergartens of which Miss M. E. Simpson is director.

All these charts are characterized by the definiteness with which the immediate aims of instruction are set up with reference to each grade. For example, in arithmetic the entire subject is divided into the following major topics:

1. Notes and enumeration.
2. Addition.
3. Subtraction.
4. Multiplication.
5. Division.
6. Tables of weights and measures.
7. Fractions.
8. Oral work.
9. Problems.
10. Decimals and percents.
11. Miscellaneous.

Each of these major topics, except the last two, is represented in the instructional work of each grade from the low-second to the high-sixth. Each of these major topics, except the last two, is represented in the chart are the grade designations. Within the compartments governed by the column and line heads are listed the subtopics pertaining to the major topic and grade in question.

Thus each teacher may know precisely what ground she is to cover in her grade. Not only this, but parents who wish information on the same question may easily obtain what they wish. This last advantage is one which is frequently underestimated. It is a lamentable fact that even the well-informed and well-disposed home usually cannot find out with any usable definiteness what the work of a given grade is supposed to be.

Accompanying the phonics chart is a pamphlet entitled "Provisional Course of Study in Phonics, Grade I–III."

We feel sure that the Rochester idea of exhibiting the objectives of teaching in the form of a conspectus of the entire course is a particularly good one. Many teachers, superintendents and principals will wish to obtain copies of these charts. We hope that the Board of Education at Rochester will print them in sufficient numbers so that they may be distributed at a nominal cost to those who can make good use of them.

The Correspondence between the Results of Illinois Intelligence and Binet-Simon Scales

As part of the admirable and astonishingly diversified research activities in Cuyahoga County (Ohio) under the general direction of County Superintendent A. G. Yawberg and the more particular direction of Assistant Superintendent Joseph A. Baer, intelligence tests have been widely used both in the centralized and one-room rural schools. A particular study was made of the correspondence between the Illinois Intelligence Scale and the Binet-Simon Scale. Dr. Bertha Luckey of Cleveland examined children with the Binet Scale. Results on the Illinois Scale were also available in 84 cases. The correlation between the I. Q.'s for these cases as obtained by these two instruments was found to be 0.67. Elementary grades three to eight inclusive were represented among these cases. The mean for the Binet I. Q.'s was 99.6 and for the Illinois I. Q.'s, 108.9.

In communicating these facts, Superintendent Baer raises the question of whether success on the Illinois Scale is not to be ascribed in part to the effect of school training. He believes such to be the case and offers the following comment in support of his position.

"It occurred to me that if the Illinois Test is affected by school training there should be an increasing difference (i.e., between the Binet and the Illinois) from grade to grade. Using the I. Q.'s for both tests I find that the third grade averages 5.7 points lower on the Illinois Scale, that the fourth grade averages 5.7 points higher on the Illinois, the fifth grade 12.7 points higher and the sixth grade 18.1 points higher. Thus far my notion is supported, but in the seventh grade the average is only 13.9 points higher, and in the eighth grade only 9.6 points higher.

"The falling off in these grades might possibly be due to the facts that the increase in the Illinois Test reaches its maximum at the sixth grade and also that the Binet-Simon tests are influenced by the same factors in the upper years. Vocabulary, reading, and general school training are increasingly prominent in the years 12, 13, and on."

We are inclined to agree with Superintendent Baer that success with the Illinois Scale is more dependent upon school training than is success with the Binet Scale. We believe, however, that a similar statement may be made with reference to all the other verbal intelligence tests.

It is generally agreed that the Binet Scale is our best instrument for measuring intelligence. Nevertheless, it too has its unreliability. Otis and

Knollin, in the September, 1921 number of the JOURNAL OF EDUCATIONAL RESEARCH showed that in terms of mental age the probable error of the Stanford Revision of the Binet Scale was 5.3 months. A part, therefore, of the lack of correspondence between the results of any group test and the results of the Binet Scale must be ascribed to the fact that the Binet Scale itself is not a perfect criterion.

Adopting the seven-fold classification suggested by Terman—genius, very superior, superior, normal, dull, border line, and feeble-minded—Superintendent Baer found that the Illinois Scale in comparison with the Binet Scale placed pupils in groups as follows:

2 pupils were placed 2 groups lower
6 pupils were placed 1 group lower
35 pupils were placed in the same group
23 pupils were placed 1 group higher
14 pupils were placed 2 groups higher
4 pupils were placed 3 groups higher

Superintendent Baer therefore points out, "if the Illinois Test can be said to err, it is in the direction of liberality, as the results for this school are consistently higher on the Illinois Test than on the Binet-Simon Test." The Illinois Scale does indeed err as all scales do; but if it must err, it is less serious that it should do so in the direction of liberality than in the opposite direction. If children are reported to be more capable than they really are, the results are likely to be less tragic both for the child and for the school than would be the case if they were represented to be duller than they really are.

The writer believes that Superintendent Baer's conclusions are substantially correct with reference to the Illinois Scale. Incidentally we may say that in the one grade (namely the third) where pupils were reported to have I. Q.'s lower than those obtained from the Binet Scale we are not in our own work recommending the use of the Illinois Scale. We believe that better results will be secured in the third grade by using some other test.

Do Teachers' Marks Measure Anything

A surprising commentary on the unreliability of teachers' ratings is afforded by an experience which a County Superintendent in Ohio has recently reported.

In connection with his quarterly examinations, Superintendent Vance duplicated two of the pupils' papers and sent them together with copies of the examination questions to 92 rural teachers. One of the papers was in fourth-grade geography and the other in eighth-grade arithmetic. The geography paper must have been a curious one, but even so, it could on no possible grounds justify the disagreement as to its value among the 92 teachers. One of them rated it 74, while three of the teachers agreed in rating it zero. Several teachers thought the pupil who wrote the paper should have 60 or more, while with the same data in hand, other teachers either thought the paper worth nothing at all, or gave it such marks as 20, 13, 16, etc. Four teachers agreed in rating this unfortunate pupil 20 for his performance in geography, but quite as many of them thought his performance was worth 50. Several teachers not content with expressing themselves with the accuracy implied by a percentage scale thought they were able to make more accurate determinations; such ratings as 43½; 48¾; and 13⅔ appear.

The eighth grade arithmetic paper—despite the supposed objectivity of arithmetic—fared even worse than the geography paper. The pupil who wrote the paper was rated all the way from 10 to 100, nor were these extreme ratings isolated from the intermediate ratings.

No fewer than seven of the teachers expressed by their ratings of 100 their opinion that the paper was a perfect performance. On the other hand two teachers agreed that the same paper was only worth 10. If 75 were regarded as the passing mark, the pupil who wrote the paper would have been passed by 47 of the teachers and failed by 45 of them. In other words, the pupil in question stood practically an even chance of passing or failing depending on the teacher who rated his paper.

Such a situation is either ludicrous or pathetic according to one's temperament or point of view. Those who doubt the value of standardized tests would do well to consider that whatever their shortcomings may be, they permit no such travesty in estimating the value of a pupil's achievement.

Salary Increases Depend upon Efficiency

We have received from Superintendent W. W. Lewton of Cicero, Illinois, a copy of his efficiency rating scale and salary schedule for 1922. The scale calls for a rating of each teacher on five factors in determination of four ability groups: (1) training; (2) health; (3) personality; (4) professional attitude; (5) success—the measurement of efficiency.

The following definition or explanation of ability groups is too concise to warrant briefing:

"Ability Groups:

ONE: Poor:—Persons whose work, or attitude, or both, have been unsatisfactory; who show little promise of growth and improvement.

TWO: Average:—(a) Those who show little, if any, improvement over the work of previous years. The results obtained may be fairly satisfactory, but the fact that a teacher has reached the point where she ceases to grow and improve, places her in the doubtful class. The work of a teacher in this group soon deteriorates and her value is greatly reduced. (b) Those whose work is uneven or not consistent; strong in some lines but weak in others which are essential. (c) Those who cannot or will not cooperate freely; are discontented and fault-finding—a menace to harmony, and a source of friction; are not positively loyal, or neglect assigned and understood duties.

THREE: Good:—Those who are strong and do uniformly good work; who measure up well in all departments and show improvement from year to year.

Those who assume responsibility in caring for supplies and records in a business-like manner.

Those who cooperate cheerfully and effectively in building responsibility and playground leadership.

Those who give careful attention to general and monthly outlines, and preparation of daily plans; and to a prompt and appreciative following of supervisors' and principals' suggestions.

Those who have considerable knowledge of, and sympathetic interest in the lives of the pupils.

FOUR: Excellent:—(Superior) Those who possess the qualifications of class Three, and in addition are distinguished in the following respects:

(1) Unusual knowledge of and interest in the lives of the pupils.
(2) Power of imaging educational problems, and offering constructive suggestions (initiative).
(3) A positive and helpful influence upon the lives of pupils (ideals and character building).
(4) The attitude of the missionary toward his work.
(5) Marked power of leadership.
(6) Extraordinary achievement in accomplishment along special departmental subjects or regular class progress."

The salary schedule is as follows: minimum $1000 for two years training beyond high school; maximum $1800; increases below $1500, $50.00 for service plus value of rating; increases above $1500, value of rating only. The values given to the rating are as follows: rate of A minus or below, dismissal; A no increase; A plus and G minus, $25.00 increase; G, $50.00; G plus, $75.00; and E, $100.00.

On the basis of service and rating ten were refused contracts; 5 received no increase; 25 received $50.00; 27 received $75.00; 51, $100.00; 10, $125.00; and 2, $150.00. These increases include both service and merit. The average increase for the 120 teachers was $83.33 and the average annual salary, $1,367.50.

An interesting experiment with retarded children is reported by District Superintendent Veit in the *Bulletin of the New York Society for the Experimental Study of Education*. About a year ago 35 pupils who were 14 years of age were taken from the fifth-year classes of neighboring schools and were placed in a special class under a very efficient teacher in the Franklin Lane Junior High School. The idea was to utilize the facilities of the school for the purpose of interesting these pupils who had found learning from books in the ordinary way a great hardship and very distasteful.

"The interest and sympathy of the teacher and her ability to make friends of this type of boy and girl, together with the making of a program that introduced sewing, cooking, typewriting shopwork, and other manual pursuits, aroused a new interest in those who in their school career were considered out-distanced.

The record of attendance for these pupils shows a remarkable change. Where formerly an attendance officer compelled their presence at school, there is now voluntary appearance. The steady progress in essential subjects observed in the work of the pupils has proved conclusively that the procedure introduced is most effective and valuable.''

At the end of the year instead of seeking employment certificates 32 of these pupils wished to continue similar work. The success of this venture points to one of the greatest possible contributions of the junior high school, that of helping duller boys and girls to realize that with proper effort and direction certain achievements are possible.

Experimental Work Spreads among Teachers of City

The *New York Evening Mail* of April 19, 1922, records the fact that scientific study of school problems is spreading among the teachers of New York. The evidence on this point is revealed in the remarkable growth of the New York Society for the Experimental Study of Education. ''Formed three years ago with a handful of members, the organization now boasts of more than 800, and it is a rare meeting which does not see 50 or 60 additional applicants admitted.'' The meetings of the society are alternately general sessions at which the entire membership discusses a common topic, and sectional meetings where the members divide into various groups for the study of individual problems. The writer of the article to which we are referring remarks that ''the Society resembles more than anything else a post-graduate school in education.''

The great size of the schools of New York; the marvelous diversity of types of instruction, and of organization; and the concentration within the metropolitan area of a great number of higher educational institutions afford an unrivaled opportunity for the study of education on a large scale. We wish well to the Society for the Experimental Study of Education.

Educational Research Association

(FRANK N. FREEMAN, *Secretary and Editor*)

THE IMPROVEMENT OF INSTRUCTION IN RURAL SCHOOLS

Communicated by W. J. Osburn,
State Director of Educational Measurements, Department of Public Instruction, Madison, Wisconsin

The statement that the use of standard tests *should* result in the improvement of instruction is so frequent in educational literature that it has become a mere commonplace. The statement that the use of such tests *has* resulted in such improvement is not so frequent. Information as to just how much improvement can be obtained is even harder to find. Such information for one-room rural schools is almost entirely lacking. The impression prevails that improvement is impossible in the type of rural schools which most of the children of the nation attend. Such an impression is disastrous in its effects and nothing could be further from the truth. Improvement can be obtained easily, quickly, and to an astonishing degree when suitable remedial devices are used. The following tables show the evidence which is now available in Wisconsin:

TABLE I. IMPROVEMENT IN SILENT READING IN WISCONSIN RURAL SCHOOLS IN 1920-21

PERCENTILES	MONROE STANDARDIZED SILENT READING TESTS COMPREHENSION											
	Grade III		Grade IV		Grade V		Grade VI		Grade VII		Grade VIII	
	B[a]	A[b]	B	A	B	A	B	A	B	A	B	A
90	10.7	16.7	16.7	24.1	24.0	27.5	28.9	34.4	30.3	39.8	36.1	44.8
75	7.6	13.2	12.6	18.6	17.2	22.1	19.6	27.5	24.5	33.5	29.5	39.6
50	4.2	9.2	8.3	13.2	12.6	17.2	13.1	20.3	17.4	26.0	23.6	31.5
Standard	8.3	10.6	13.3	15.6	17.2	20.5	18.1	20.8	26.0	27.3	28.2	29.4
25	2.1	6.3	4.5	9.0	8.6	12.4	9.1	15.0	11.7	20.0	18.6	24.9
10	0.8	3.6	1.8	5.8	4.8	7.9	5.3	10.6	6.4	13.2	11.7	18.4
Number	149	149	200	200	184	184	200	200	136	136	171	171
	Grand Total											1040

[a] B, scores before remedial instruction.
[b] A, scores after remedial instruction.
Time intervening—four to five months.

TABLE II. IMPROVEMENT IN READING ABILITY IN ONE-ROOM RURAL SCHOOLS

PERCENTILES	GRADE II		GRADE III		GRADE IV		GRADE V		GRADE VI		GRADE VII		GRADE VIII	
	B[b]	A[c]	B	A	B	A	B	A	B	A	B	A	B	A
90........	126.7	134.8	120	[a]120.7	119	[a]123.1	124.8	[a]121.9	117.9	[a]121.5	108.5	[a]119.4	109.8	125.3
75........	121	122.3	108.4	118.4	108.7	117.1	108.8	119.9	107.3	118.5	101.1	113.8	105.8	115.4
50........	101	108.5	93.75	107.1	96.0	105.0	96.5	109.8	94.2	103.5	92.3	102.1	93.7	106.0
Standard..	100	100	100	100	100	100	100	100	100	100	100	100	100	100
25........	90.0	93.5	82.6	95.3	85.0	94.3	85.4	96.9	82.6	94.3	81.7	91.8	85.3	93.0
10........	76.4	78.4	74.5	85.8	75.2	81.5	80.1	87.5	76.1	85.7	73.1	80.1	77.0	87.1
Number...	114	68	109	103	121	115	139	130	131	134	141	123	139	112

[a] Not reliable. Quotients from one county not given in detail above 120.
[b] B, scores before remedial instruction.
[c] A, scores after remedial instruction.

Table II shows the improvement in reading ability after an interval of four or five months has elapsed since the original testing, during which time remedial instruction was given. A different form of the test was used in each case. The testing was done by trained supervisors. Seven hundred and eighty-five children took the tests, upon which the conclusion is based.

Two observations can safely be made concerning the foregoing tables:

1. One-room rural schools can be quickly and easily brought up to standard by intelligent supervision and the use of up-to-date methods of teaching.

2. The amount of improvement is three or four times as great as that which is usually obtained in the same length of time. In Table II the average gain is 10 percent in reading quotients. This is equivalent to the gain of a year. And this gain is *over* and *above* that which is ordinarily expected since reading quotients are supposed to remain fairly constant from year to year.

JOURNAL *of* EDUCATIONAL RESEARCH

Volume VI NOVEMBER, 1922 Number 4

TECHNIQUE OF THE PUBLIC SCHOOL SURVEY[1]

J. B. SEARS
Stanford University

DEVELOPMENT OF THE SURVEY MOVEMENT

Origin of the survey idea.—The survey movement in education began in 1910 and 1911 with surveys of the school systems of Boise, Montclair, Baltimore, East Orange, and Boston. For several years previous to this date cities and states had been subjecting their schools to the careful scrutiny of lay commissions, while in our universities a quantitative terminology was slowly being introduced into educational studies. These lay investigations and scientific studies were the immediate precursors of the educational survey movement.

The forces which had brought these several activities into existence were the same forces that had brought the business expert or the efficiency engineer into business and industry. The taxpayer had begun to ask, "what do I get for the dollar I have invested in schools?" People everywhere had begun to criticise the management of the schools, the curricula offered, and even the aims toward which the schools were working. Meantime students of education were at work on the development of a technique for the scientific study of educational problems.

It may be shown that the survey movement has made two important contributions toward the development of a science of education. One is in the extensive and clear analysis it has made of educational facts and situations. The other is in the additions it has made to the actual methods and technique of educational investigation.

---

[1] This paper was presented at the Conference on Educational Research and Guidance held at State Teachers College, San Jose, California, May 19 and 20, 1922.

To substantiate these statements one only has to run through the survey reports. Nowhere before this do we find such clear and such well-reasoned statements of educational aims. And the same may be said with reference to the functions and character of buildings, curricula, supervision of health, finance, organization, and administration as well as of many lesser topics. On the side of technique the survey movement has made substantial contributions to tests and measurements, having worked out and demonstrated the use of quantitative standards in a wide variety of places in education, and having brought to light a goodly number of working principles with reference to which teaching, administrative, and supervisory processes may be directed. Besides these specific contributions it has furnished a most valuable general stimulus to scientific work in education.

Dominating purposes in the survey movement.—After a little more than a decade of development and of service to the public we may properly ask, what have we learned about surveying? The work has succeeded in the sense of having convinced the public of its practical worth, and it is so bound up with the movement which is devoted to the scientific study of education, both in its workers and in the purposes and character of its work, that its value in this direction is also obvious.

As we look over the brief history of the survey movement, it is possible to distinguish three main purposes which have dominated the work almost from the beginning. The first and foremost of these purposes has been a purely practical and utilitarian one, viz., that of helping the schools to a better practice. Back of this purpose were the demands of the people for an accounting and the earnest desire of our leading educators to bring the schools of the country up to better standards.

The men who have really directed the survey movement have been scientific men, men who have not only conceived the highest achievement of any science to be its ultimate issue in benefits to mankind, but who have realized that such benefits can come only in proportion to the scientific results that are achieved. To make education scientific, therefore, was also their ambition, and in the school survey they saw the opportunity to learn new things about education and to apply, and perhaps discover, new methods 2 of study. A second purpose, therefore, has been the gathering and analysis of new facts, or the increase of our knowledge 3 about educational phenomena; and a third purpose was to apply and to develop the scientific technique essential to the building up of a science of education.

Without these subsidiary purposes it is doubtful whether the best educational leadership could have been so largely drawn into the movement as it has been. If then, some of the reports of surveys have been over-technical in form, we should feel no particular surprise, and perhaps not too much regret, since it has meant at most only a wrong emphasis and a more indirect attack upon the practical problem and, to some extent, these have been offset by a relatively larger scientific contribution.

Development of the proper functions of a survey.—What are some of the fundamental facts and principles with reference to which the survey movement and the individual survey are to be directed? What has "survey" come to mean; what may the survey movement hope to accomplish; and what particular principles and technique have been developed for this type of work?

Very early in our survey experience the question arose as to how far a survey should go in uncovering graft, dishonesty, and unfairness, as distinct from other types of inefficiency. Should a survey subpoena witnesses, conduct hearings, and try to place the blame for inefficiency or crookedness upon the particular persons found responsible, or should it stop with a clear statement and explanation of the inefficiency, leaving to the board of education or other properly constituted authorities the task of placing the blame upon persons or of bringing the matter to the attention of a grand jury?

In spite of opportunities no important survey assumed the responsibility for exposing graft or for referring by name to those responsible for inefficient service. It was early declared, and the position has consistently been maintained, that such work was the function of grand jury or other special investigation, and not the function of a survey. This is a significant distinction since it has removed from the realm of the survey a large amount of work of a purely negative type and forced upon surveyors the necessity of finding a positive and constructive function to perform.

A second question that had to be settled was, how far shall the survey go in outlining a program of future action for the schools. Was it enough to describe and explain what was good and what was poor about the schools, or should the survey show how to do away with what was poor and how to replace it with something better? To settle a question like this was not easy, and there appears to have been difference of opinion. Many of the early reports, and not a few since then, have been quite largely descriptive in character with special emphasis upon the weak points of the system in question. But this is not strange when we consider how little we had in the way of standards or even of authoritative opinion on what a city school system should be like.

At the beginning of the survey movement, studies of retardation were still in an undeveloped stage, pedagogical tests were just being made, but were hardly available for use in a survey, there were no group mental tests, and practically no standards available for use in the study of buildings and equipment. Not much was available as a basis for judging staff, curricula, organization, supervision, and administration; and only a little had been done in the comparative study of school costs. Without standards, without an adequate knowledge of practice over the country in any of these major fields, it was not easy to outline a constructive program for the future.

In the face of these difficulties, however, it soon became evident that a survey must project into the future if it was to be useful. In the light of our experience to date, then, we may define the functions of a survey about as follows: The function of a school survey is to apply the best available technique in the study of a school system, with a view to revealing the character and the extent of its efficiency, and with a view to showing how the best possible results may be obtained. This calls for analysis and measurement, for the use of exact methods, for the application of principles and quantitative standards wherever possible. It calls for description and for interpretations, for a statement of facts, whether they are favorable or unfavorable, and for constructive proposals as to what ought to be done next and how to do it.

Early lack of survey technique.—Having settled upon a proper meaning, or function, the survey turned naturally to the question of how to accomplish its purpose. This immediately raised the question of how extensive a survey must be in order to afford an adequate basis for outlining a future program. It also raised the question of whether a single feature of a system could be surveyed separately. Practice soon answered this latter question in the affirmative. Numerous studies of "progress of children through the grades," of "curricula," of the "efficiency of pupils in their studies," and of "finance," were undertaken as surveys of these aspects of the schools.

But how to measure the efficiency of a school system as a whole was a difficult question. When the technique for studying retardation had reached a fairly satisfactory stage, retardation studies were heralded as real measures of the efficiency of school systems. One has only to read the early literature on retardation to be convinced that the development of this one piece of technique was a very striking thing in the recent history of education in this country. He may also be struck by the credulity with which the rather crude interpretations of the results of retardation studies were accepted.[2]

At the outset there were almost no instruments for making exact measures of the processes, machinery, or products of education; and no great skill had been developed in the use of comparative methods. The practical demands of the survey disclosed the fact that what education had in the way of facts, methods, and technique had been developed and, in the main, adapted for use in the laboratory and classroom. We had worked largely in isolation, and with reference to the demanc of logic. Now we were asked to work in the open field, in tl midst of the tangle of practical conditions, and with referenc not only to logic, but also to "what will actually work." It wa clear, therefore, that the survey had before it the task of buildin up a technique of its own.

---

[2] See: Volkman, Hilda, and Noble, Isabel, "Retardation as indicated by one hundred city school reports." *Psychological Clinic*, 8:75-81, May, 1915. Presents a study of the history of the influence of retardation studies on city superintendents' reports.

SURVEY TECHNIQUE

The meaning of technique.—By survey technique we mea the particular methods and machinery essential for the carryin on of a survey. It includes all the various instruments an devices used to detect the presence or absence of a thing, or t determine the amount of the thing present, or to determir causal or other relationships between things found. It includ principles, axioms, plans, and opinions. It also includes tl methods by means of which these instruments and devices, the principles, axioms, plans, and opinions, are brought to bear i determining the aim, the extent, and the plan of the survey, well as in carrying forward the process from its inception t its final conclusions and recommendations.

It is not our purpose to trace the history of the developme of this technique, but rather to discuss briefly some of its mo important aspects as they have thus far been built up. The we shall consider as nearly as possible in the order of the ne for them in making an actual survey.

Defining the aim of a survey.—First of all, the survey mu have a definite aim. This seems a very trite statement, yet one has read the survey literature, one will realize its pertinen here. Obviously the function which the finished survey is e pected to perform must provide the aim which directs the su veyor in his work. Yet, as our survey reports show, it is ve easy for the surveyor to be overwhelmed by the academic scientific possibilities which at once appear in the study of school system. Accordingly we should set down as a workir principle to be kept in mind that *the aim of a survey must determined by the functions which are finally to be served by t* *completed task.* We should also add, that the plan of procedure in the study must be drawn with careful reference to these aims.

The relation of the aim to the plan of work.—The relation of aim to working plan suggests the second problem. The difficulty here lies in the fact that we must determine what facts, processes, and conditions to study, what principles, methods, devices, and instruments to apply in the attainment of the aim, and how to organize our procedure with respect to a plan of operation. The broader and more comprehensive the aim, naturally, the more difficult it is to devise a plan of procedure that will unravel the many complicated relationships with which the survey must deal.

To the writer there seem to be four main steps or problems in any complete study of a school system. First, the community's educational needs must be defined. Second, the community's capacity to provide for its needs must be ascertained. Third, the present educational facilities must be evaluated in the light of what is needed and in the light of what the community is able to pay for. Fourth, a constructive program of procedure must be worked out with reference to such of the corrections and expansions recommended as the community is able to provide.

In a sense these are four distinct pieces of investigation, each of which must be carried out in any complete survey. It would be possible to carry them out separately and in the order here presented. No survey, however, has done this; and the reason for this is clear. To locate some of a community's needs one would of necessity make many studies almost exactly like the studies necessary for measuring the efficiency of schools as they exist. Furthermore, in reporting the survey such a plan would involve repetition at so many points that it would not be economical. The point to emphasize here is that these are four functions to be performed in a survey. They suggest not only the main divisions of the aim of the survey, but also the logic of the plan of work. In the latter sense they stand as a constant check on every step in the study. If the above principle for selecting the aim of a survey is our first piece of technique then the plan here suggested is our second piece.

What is the technique essential to the carrying out of these four lines of study? Since the character of the study determines the kind of technique, we shall undertake to sketch out in very brief outline a plan by means of which each of these four ends or functions may be attained. By this method the technique will be discussed in connection with the function it is to perform in making a survey.

Problem one: How to determine a community's educational needs.—In considering, first, the technique for determining a community's needs for education we may lay it down as almost axiomatic that *a community's educational needs are determined by the character of its own culture.* The full significance of this axiom will not readily occur to any except scientifically trained educators. Consequently its use in surveying is a much more technical matter than is its mere recognition as a self-evident truth. A people's economic, political, social, aesthetic, moral, physical, and intellectual interests, ideals, and activities go on together. Change in any one of these will produce some change in all the others. If a high type of culture is the aim then what we do in education must have reference to all the elements that together constitute what we call culture or civilization. It is in the light of possible corrections, expansion, and enrichment in all these lines that educational needs must be defined.

Educational needs must be considered from the standpoint of extent, of kind, of quality. We need to know how many buildings, how many teachers, how many courses; we need to know what kind of buildings, courses, equipment, organizations; and we need to know what degree of excellence in all these things it is best to provide. What facts about the community do we need, what instruments of measure may we apply and how shall we interpret our findings?

If education is to have respect to the social, economic, physical, aesthetic, moral, and intellectual life of the people, then our facts will be drawn from these fields. If the application of the above axiom in the determination of needs is step number one in our technique, then step number two would be the selection of a proper basis for judging what facts to select. Since our needs are both quantitative and qualitative in education, we shall need to reveal the quantitative and qualitative aspects of these many features of our culture if we are to set up a complete picture of our needs.

Without pretending to present the full line of reasoning upon which a selection should be based, I shall merely attempt to outline briefly the facts that are needed, to indicate the sources from which they may be drawn, and then to explain how these facts are to be treated.

MATERIALS AND SOURCES FOR A SURVEY OF A COMMUNITY'S EDUCATIONAL NEEDS

I. Data needed:
    1. Carefully drawn maps showing:
        a. Size of district.
        b. Main features of its physical and industrial geography.
        c. Location of public buildings, parks, and playgrounds.
        d. Natural population centers.
    2. Information about the population.
        a. Extent of population through a period of years.
        b. Racial composition.
        c. Age and sex distribution.
    3. Information about literacy.
    4. Information about occupations.
        a. Number and kinds.
        b. Workers in each.
        c. Sex and age of workers.
    5. Public enterprises.
        a. Government.
        b. Industrial or commercial activities.
        c. Provisions for social, intellectual, and aesthetic interests.

II. Sources of information:
    1. National and state census reports.
    2. Municipal directories.
    3. Official records in offices of city auditor and engineer.
    4. Chamber of commerce literature.
    5. Educational reports and records.
    6. Historical materials.

Our purpose is to arrange these facts so as to present a clear picture of what a school system should do to correct weaknesses or to expand and enrich the various aspects of community life. By the maps we can detect natural attendance barriers, such as streams, bad crossings, and heavy traffic centers. We can thus show any existing need for reorganization or extension of attendance district boundaries. We can also see where to locate new buildings in order to utilize parks, playgrounds, and museums in connection with instruction. From our population data we can see about the rate at which school facilities must be expanded to meet future needs; and from the same source we may detect certain special curriculum needs, such as the adult education required by foreign groups. We can see by our age distribution the extent of the school population. By our literacy facts we can see the need for extension, Americanization, and adult education. From our occupational data we can see clearly the economic background with respect to which attendance, special courses, types of special schools, and indeed the entire curriculum and management of the schools may be worked out. Information as to the kinds of occupations and the extent to which they are available for boys or for girls is basic as an index to educational aims. Facts about the local government in connection with facts about race and industry are suggestive with reference to the teaching of history and civics, and with reference to the general social life of the schools. Municipal industries ought to be used as educational laboratories for the youth who are to direct these enterprises a few years later. Facts about the city's interest in beautiful streets, buildings, and parks, in music, libraries, and museums afford a basis for defining the community's needs for training along aesthetic and public-enterprise lines. These are but the broadest lines of school policy. By pursuing our study in some detail, community by community, within a city or county, we may arrive finally at the size and character of buildings and of classrooms, the character and extent of equipment for playground, library, laboratory and classroom, the number and kind of teachers, and so on to the end of greater and lesser needs for amount and kind of education.

The above are but samples of information, or hasty views, from the rich picture of educational needs that may be obtained from these facts when they are properly organized. Some of these are facts as to quantity or extent of educational facilities needed, while others reveal the particular character or quality of these facilities. The basis of our choice of data may be defined by the following principle: *Educational need depends upon such facts about a community and its life as will reveal the extent, condition, character, and tendencies of its social, political, economic, intellectual, aesthetic, and physical life.*

A third step in our technique has to do with the instruments, methods, and devices to be used in assembling and classifying these data. This step is relatively simple. It calls for tabular and graphic presentation of facts in isolation and in such serial and comparative manner as will reveal changes associated with the passing of time, and the size or extent of the phenomena in question in comparison with like facts in comparable communities elsewhere.

This technique is the most exact we have for measuring such factors as we are discussing at this point. We have no precise formulae for prescribing the number of new educational units (units of buildings, units of equipment, units of curriculum) called for by the addition of a unit of population to a community. We must therefore base our judgment of need upon rough common sense estimates and upon what is common elsewhere under like conditions. If this is a crude technique, it is at least trusted in business, industry, and politics; and it must serve here until a finer procedure is worked out.

A fourth bit of technique is the interpretation of the facts brought out by our tables. Here we are called upon to make an educational prescription with reference to a certain set of social or economic facts about a given community. What we do is to consider how much of a certain kind of educational activity is necessary to produce or modify in character or amount the particular social or economic phenomena for which our facts stand. This consists in the application of certain accepted educational principles and standards to a defined situation. To say here what principles and standards is obviously not possible except by way of illustration.

Let us say that our facts show in one part of a city a large industrial community where the people are mainly foreigners, and where large numbers of young people of both sexes are employed in factories at as early an age as possible. How should we express the educational need which these facts reveal? What do these facts say about the subjects of civics and mathematics? Clearly the foreign child is less familiar with American social and political ideals and practices than are native children. Accordingly, the need for civics instruction in this community is relatively large. Add to this the fact that the children leave school early and go to work in factories. This means a further increase in need for civics teaching, and it also means that the mathematics should be connected up in a practical way with the points in industry with which the children are soon to associate themselves. Plainly, too, it argues for breadth in the information and experience to be brought into these, as into all courses, if we are to develop among these foreigners a culture that will best fit them into our American scheme of civilization.

Here we selected the principle that, *every child who is to become a citizen should be trained in the fundamentals of citizenship,* and the further principle that, *education must furnish a child with concrete contacts with the economic processes in which he will participate as an adult,* and roughly speaking yet another principle that, *educational need for a given kind of training exists in inverse proportion to the amount a child possesses by virtue of his social inheritance plus his experience and training to date.* Our technique of interpretation is the technique of educational diagnosis and prescription, and consists in selecting the principle or standard that applies and, by referring it to the facts in question, judging the amount or kind of education necessary to originate or modify those facts.

It is not necessary to say that this is not a routine matter, not a technique that can be learned in a few lessons. Nor should it be necessary to say that in applying these four bits of technique we are bound by the same logic that directs scientific study anywhere. Indeed, the facts we have proposed to assemble for the purpose of studying a community's educational needs might be assembled in a somewhat similar manner for other purposes. It is not a new logic, or a new set of principles of research that the survey has built up, but rather a particular application of old logic and old principles to new ends. It is only in this sense that the technique here discussed is new.

✓ *Problem two: How to determine a community's capacity to meet its needs.*—Here, as in the first of our four main survey problems, we shall present a practical outline for conducting the part of the survey having to do with the community's ability to meet its educational needs. Our first task, however, is to decide upon what basis we may say that a city is or is not able to support a given plan of education. Here again we must resort to the use of general social and educational principles. The first principle needed is: *A democratic state, being dependent upon the intelligence of its citizens, is obligated to provide some amount of education for every child.* A second principle is: *Possession of the simple tools of knowledge and some understanding of our social, economic, and political past and present, being essential to good citizenship, must be provided if the state is to have a proper margin of safety.* A third principle should be: *Beyond the minimum provision, and, up to full satisfaction of need, the amount of education provided should be determined by a community's financial ability.*

If we accept these principles and attempt to apply them, we must find a way to measure a community's financial capacity. In the following outline of this feature of a survey, we shall show the kinds of facts needed and where these facts may be obtained. We shall then explain how they should be treated. The decision as to what facts are necessary rests upon the simple truisms, (1) that a community's wealth or income determines the limit of its expenditures, and (2) that education is but one of several activities that must be paid for from the surplus of income over necessary expenditures. First, then, we need to know, how much wealth and earning power a city has, whether these are stable or fluctuating, and, if fluctuating, how and at what rate changes take place. Second, we need to know what the necessary expenditures are.

MATERIAL AND SOURCES FOR A SURVEY OF A COMMUNITY'S CAPACITY TO MEET ITS EDUCATIONAL NEEDS

I. Data needed:
   1. Population with rate of increase or decrease.
   2. Percent of dependents in population.
      a. School age and younger.
      b. Those retired from active life.
      c. Physical, mental, and moral dependents.
   3. Taxable and real wealth of community.
   4. Tax rate, total and for different purposes.
   5. Income or earning power of community.
   6. Dependability of sources of wealth.
   7. Amount and distribution of all public expenditures.
   8. Rate of increase in wealth and tax rate through a period of years.

II. Sources of information:
   1. Federal census reports of finance, population, and industries.
   2. Auditors' reports.
   3. Tax assessors' records.
   4. Local, state, and federal school reports.
   5. Income tax records.

It is impossible to present here full reasons for the inclusion of all these data. The above principles must serve as our general justification. Briefly, however, a study of real and taxable wealth will indicate the extent of the sources of income. The percent of dependents in the community will enable us to estimate the number of aggressive earners of wealth. A knowledge of the rate of change in population and wealth will permit us to anticipate possible income in the future. We need to know tax rate in order to see how heavily the community's income is now being drawn upon. We need to know the details of public expenditures because education is but one public interest calling for outlay. Finally, we need to know the tendencies in all these things, and we need to know them in comparison with similar communities elsewhere. All these we need to know for reasons that are reasonably obvious when their relationships are apprehended.

Our task, then, is to assemble the data to meet these ends, and to find some basis for interpreting the display when it is arranged. Tabular and graphic representations, temporal series of facts, and tables of comparative data are what we need in order to show exactly what a community has in the way of purchasing power, and how it has been using its purchasing power. The technique of the selection and arrangement of data, and of determining a basis of measures, are no different from that discussed above. Here, as before, we measure a thing against itself, and against others of its kind. Experience is the basis of our judgment and the only basis we have so far devised. It is largely a matter of careful inspection and comparison.

Interpretation calls for judgment as to the importance of education as one of several public enterprises and as to when a community is spending up to its reasonable limit. It calls for judgment as to the stability of public income. In these processes we apply the comparative method, and such general principles as, that education should ordinarily call for a large expenditure in comparison (for example) with police protection, and that every source of wealth should be made to bear its share of the public burdens. But to go into the details of the reasoning necessary to full interpretation is impossible here.

Problem three: How to evaluate the present efficiency of the schools.—So far it appears that the survey method is a sort of rough and ready reasoning. Thus far we see an array of facts, and analysis of the facts, and the application of axioms and principles but no exact measures of anything. This only serves to point out the difficulty of survey work. It does not argue that there is no special survey technique.

In handling our third function in a survey, that having to do with the evaluation of the present efficiency of a school system, we shall introduce, in addition to the methods thus far described, a more exact procedure. Here we are concerned with a measure of things as they are. How much education have we; how well organized and managed is it; what is the value of the training given; how effective is the training; how adequate the housing and equipment, the staff, the supervision, and the curriculum? These are the types of information we now seek. To show the nature of the procedure, let us first examine an outline of the information needed, together with the sources from which such information can be secured.

MATERIAL AND SOURCES FOR A SURVEY OF THE PRESENT EFFICIENCY OF A SCHOOL SYSTEM

I. Data needed:
   1. The larger problems of administration.
      a. The present school policy.
      b. The legal and administrative organization.
      c. Finance and business management.
      d. Buildings and the building program.
      e. The teaching and supervisory staff.
      f. The system of records and reports.
   2. The internal administration and management.
      a. Child accounting.
      b. The course of study.
      c. Grading and promotion.
      d. The daily program.
      e. Vocational and educational guidance.
      f. The system of grades and marks.
      g. Equipment and supplies.
   3. Efficiency of instruction and supervision.
      a. Instruction measured.
      b. The supervision of instruction.
      c. Health supervision.
      d. The place and function of research, investigation, and experiment in a public school system.

II. Sources of materials:
   1. Practically all the sources mentioned above.
   2. All school records, reports, laws, and regulations.
   3. The schools and community as they can be studied by inspection and measurement.

It is by no means possible to discuss here all the methods of investigation suggested by this outline, nor even to enumerate all the different kinds that may be needed. Such methods as have been discussed above apply at many points. In addition we have a special technique for studying the rate of progress children are making in their work, another for measuring the efficiency of buildings, another for instruction, another for grading and promotion, another for the daily program, another for class marks, another for the staff, another for finance, and so on. Moreover, the findings of one of these studies may be brought to bear upon another (for instance, the results of the study of retardation as a basis for judging the administration of the curriculum).

In the development of these various methods we have not only had to analyze the problems so that we could locate and define causal relationships, but we have had to figure out ingenious devices for finding the particular truth we need in a given case; and finally we have sought some norm or standard against which to measure the product brought to light by the application of our device. Everywhere we have had to decide what information is needed, what combinations of facts will give this information, and how these combinations can be made. Axioms, principles, devices, or standards must be selected accordingly as the one or the other applies. The most technical task is not to obtain a knowledge of these tools, but to know when, where, and how to use them to get at the truth.

By way of illustration suppose that the efficiency of instruction is to be measured. Obviously our first need is an instrument of measurement. Let us use a standard test or scale. The selection of a test or scale then becomes a question. Clearly we want the one that will most accurately measure the thing in question. But what is this thing? Here, among other things, we must choose between measuring what the schools have taught (or intended to teach) and what we think ought to have been taught. Our aim in measuring will have a direct bearing upon our interpretation of results. After we have our instrument and have decided what to measure, we must then carry out the various processes of measuring. We must apply the test, assemble and interpret the scores and define the results. Each of these steps calls for some bit of technical knowledge or skill. When we consider the wide diversity of processes, products, and materials a measure of which our outline calls for, we see the need, not for a brief paper, but for a large volume on the technique of surveying. This one brief illustration can at best suggest only vaguely the many special knowledges and skills which a surveyor must have at his command for this part of a survey.

Problem four: How to outline a constructive program for the future.—Finally, in a survey, we come to the end of investigation. We have before us the facts as to what the community needs, as to what it can afford, and as to what it is getting in the way of school efficiency. The next step is to say what all this means in terms of a future program. Here we must make clear where the aims, the machinery, and the processes of the schools are weak or short of what the community is able to have, and then show what needs to be done to bring the schools up to their best possibilities.

When we show that a piece of organization, a process or an aim is weak or wrong, we must use a suitable terminology. Similarly, when we point the way toward the amount or kind of aim, organization, or process which the community should have, we must have a basis or standard with reference to which this advice is given. These standards are quantitative if we refer to success in studies or to amount of progress in grades; but if we refer to aims and purposes—and often when we refer to processes and products—we must utilize axioms, general principles, and experience elsewhere as our bases or standards.

For instance, if we wish to advise the schools to spend more money on supervision, we must show by test results, by comparison of the amount of supervision provided with the amount provided elsewhere, and by an analysis of the supervisory organization that poorer results are obtained than elsewhere, and that wrong principles of organization are in use. We must also show what results can be secured, what principles of organization should operate, and what amount of money should be spent. Expenditures in the community in question must be compared with those for the same purposes in comparable communities elsewhere. Results likewise must be compared. The conviction may finally be created, through a process of deductive reasoning, that the things recommended—things proved to be both desirable and possible—ought to be put into effect in the school system in question.

Conclusions.—It has been the purpose of this paper to bring to light something of the school survey technique. Incident to the treatment it has been necessary to discuss the nature and purpose of a survey and to offer in outline a survey plan.

As presented the technique includes a choosing of aims and facts and the formulation of plans of study. It calls for the selection and application of axioms, principles, devices, standards, and comparisons in the examination of the aims, facts, relationships, and conclusions dealt with. In the main, the technique is new and specific only in the sense that old principles, old logic, old standards are applied to new ends, under new conditions, and in new combinations and relationships. In range the survey technique includes the simplest processes of counting and the most exacting processes of reasoning. It includes the application of simple facts or truisms as well as the most complicated statistical devices.

The mastery of such a technique can come only with years of study. We have made great progress so far, but the work yet to be done is enormous. We must have more definite measures, more clearly defined principles, more short method devices. These will come partly as a result of laboratory work and partly from hard experience gained by survey work. The field is rich with possibilities of great practical service and great scientific achievement. In a sense the survey has provided the greatest stimulus that the scientific study of education has yet received because it has furnished the highest goal of scientific study, namely, the opportunity to render needed service.

MINIMUM ESSENTIALS IN PLACE GEOGRAPHY

P. R. STEVENSON
Bureau of Educational Research, Ohio State University

THE IMPORTANCE OF A MINIMUM LIST

Modern methods of teaching geography emphasize human relationships and the application of geographic principles to everyday life. One of the essential outcomes, however, should be a knowledge of important places. Indeed, such a knowledge is necessary for the appreciation of the activities of people and in making proper adjustments to environment.

The large number of places mentioned in the daily newspapers illustrates the need for knowledge of place geography. In an examination of the first page of the *Columbus Despatch*, October 1, 1922, the author found a surprisingly large number of places indicated. Table I shows these places and the number of times they occur.

A similar study of the first ten pages of the *Literary Digest*, September 30, 1922, which contains a weekly digest of important newspapers, showed the names of the following places occurring five or more times: Smyrna, Constantinople, New York, Wash-

TABLE I. FREQUENCY OF MENTION OF GEOGRAPHICAL PLACES ON A NEWSPAPER PAGE

PLACE	Number of Times Mentioned	PLACE	Number of Times Mentioned
Constantinople	11	Erin Keni	1
Ohio	8	Rome	1
Washington, D. C.	5	Great Britain	1
Paris	5	Dardanelles	1
Cleveland	4	Amsterdam	1
Toledo	4	Doorn	1
Athens (Greece)	3	France	1
United States	3	Massachusetts	1
London	3	Columbus	1
Chanak	2	Italy	1
Smyrna	2	Mudania	1
San Francisco	2	Kara Bournow	1
New York City	2	Chicago	1
Panama Canal	1	California	1
Dayton	1	North America	1

ington, Philadelphia, Europe, Turkey, France, Great Britain, and the United States. Forty-four other names of places occurred one or more times.

Pupils cannot be expected to remember all the places mentioned or discussed in class. Nor can any course of study anticipate the large number of places which current papers or magazines may mention from time to time. As a matter of fact, it is doubtful if the majority of pupils in American schools even know where the world's most important places are located. In June, 1920, Dr. B. R. Buckingham and the author made an investigation of the ability of pupils to locate important places both in the United States and in other parts of the world. A large number of place-geography items were arranged in the form of tests and submitted to pupils. The items on United States place geography were given to 6,339 grammar-grade pupils in the following states: Illinois, New York, Florida, Oregon, Arizona, Connecticut, and Montana. The items on world place geography were given to 7,849 grammar-grade pupils in the following states: Illinois, Maine, Connecticut, Colorado, and Kansas.

The results of this investigation showed conclusively that the average elementary pupil's knowledge of place geography is very limited. Table II is a summary of the results obtained on a few of the items. For example, it will be noted that only 65 percent of the pupils knew that Egypt is in Africa; 66 percent, that Persia is in Asia; 44 percent, that the Himalaya Mountains are in Asia; and so on. The pupils' knowledge of their own country seems to be no better as indicated by the last questions in the table.

The examination of such data leads one to ask the question, "Why do pupils know so little concerning the location of places in their own country, and the world as a whole?" In the author's opinion, this may be explained by the fact that among the large number of places which are presented to the child in his study and recitations, no definite distinction is made between the places of major and minor importance. The pupil, in attempting to cover too much ground, fails to fix either the important or unimportant places in his mind. If this explanation is correct, he

TABLE II. ACCURACY OF RESPONSE ON SELECTED ITEMS

QUESTIONS	Percent of pupils in grades V-VIII who answered the questions correctly.
Egypt is on what continent?	65
Persia is on what continent?	66
The Himalaya Mountains are on what continent?	44
The Andes Mountains are on what continent?	59
The Baltic Sea is nearest to what continent?	69
The Nile River is on what continent?	62
Into what ocean does the St. Lawrence River flow?	80
Into what ocean does the Amazon River flow?	70
Armenia is on what continent?	27
In what state is San Francisco located?	80
In what state is New Orleans located?	55
In what state is Cincinnati located?	40
In what state is Louisville located?	27
In what state is Omaha located?	23
In what state is the Garden of the Gods located?	15
In what state is Mt. Shasta located?	17
What state is next north of Kansas?	38
What state is next west of Connecticut?	25
What state is next east of Oregon?	69
What state is next south of Kentucky?	38

should have some drill on the location of a minimum number of important places. Miss Day has mentioned this fact in her excellent article on the Teaching of Place Geography.[1] She says:

"The first problem that confronts the teacher is to decide what places shall be taught. Place geography being only a phase of the larger field of the subject, it is necessary to guard against the teaching of too many places. A list of places to be used as a basis for the work may be provided by the teacher, or by using lists found in school journals, various courses of study, maps, and other geographical publications."

THE SELECTION OF A MINIMUM LIST

The average grade teacher does not have the proper background for selecting the most important places. A carefully selected list of prominent countries has been made by Branom.[2]

---

[1] Day, Rachel E., "The teaching of place geography." *The Journal of Geography*, 21: pp. 109-115. March, 1922.

[2] Branom, M. E. and F. K. *The teaching of geography.* Ginn and Co., 1921, pp. 276-277.

In determining this minimum list of countries Branom used the following criteria: (1) the area of the country in square miles; (2) population; (3) total value of imports; (4) total value of the imports from the United States; (5) total value of exports to the United States; and (6) total value of exports. Concerning these criteria Branom[3] says: "These criteria not only offer objective data of great significance in the study of a particular country but also afford a means of measuring the relationships of the United States to each country." A list of fifty-two countries was thus set up and ranked in order of importance. The United States, Great Britain, and Germany headed the list, while those of lowest rank were Salvador, Paraguay, and Siberia.

To obtain a list of other important place-geography items, the Bureau of Educational Research at Ohio State University submitted a large list of places, rivers, and other factors to fifty-five geography experts. These people—heads of geography departments in normal schools and universities, and supervisors of geography in city school systems—were asked to cross out any item which they did not consider of sufficient importance for the pupils to learn and remember. As a result of the judgments of these fifty-five people, five lists were made up. The first list consisted of items which, in the opinion of all the judges, elementary-school pupils should learn and remember; the second of those upon which 90 to 99 percent of the judges agreed; the third list of those upon which 80 to 89 percent agreed; the fourth list of those upon which 70 to 79 percent agreed; and the fifth list of those upon which less than 70 percent agreed. The following lists are examples of the items found in the different groups.

TESTING MINIMUM ESSENTIALS

A complete list of the places which 90 percent or more of the judges considered important were embodied in tests.[4] The United States Place Geography Test is composed of three forms

---

[3] *Op. cit.* pp. 273.

[4] Place Geography Tests. The World and the United States, Forms I, II, and III, by B. R. Buckingham and P. R. Stevenson, Bureau of Educational Research, Ohio State University. Published by the Public School Publishing Company, Bloomington, Illinois. 1922.

TABLE III. THE PERCENT OF JUDGES CONSIDERING CERTAIN SELECTED PLACE-GEOGRAPHY ITEMS IMPORTANT

List 1 (100 percent)	List 2 (90–99 percent)	List 3 (80–89 percent)	List 4 (70–79 percent)	List 5 (Less than 70 percent)
Gulf of Mexico	Persian Gulf	Baffin Bay	Mohave Desert	Cape Cambodia
Cape Horn	Cape Cod	Parana River	Amur River	Blanco Bay
Ganges River	Yukon River	Lake Baikal	Cyprus	Lake Peipus
Lake Superior	Lake Winnipeg	Tasmania	Salvador	Dvina River
China	Irish Sea	Albania	Foochow	Gulf of Oman
Alaska	Formosa	Mt. Hood	Keif	Mt. Logan
Mt. Everest	Niger River	White Sea	Perth	Lake Ladoga
Berlin	Bombay	Aral Sea	Lake Chad	Mt. Kenya
Rome	Amsterdam	Honduras	Cape Verde	Soochow
Seattle	Pyrenees Mts.	Tunis	Murray River	Fusan

with 60 exercises each, and the World Place Geography Test of three forms of 70 exercises each. One of the distinguishing features of these tests is that the pupils do not need individual printed blanks. One printed booklet for each teacher or school is all that is needed. The pupils write their answers on any ruled paper which happens to be used in the school. The cost of giving the test is reduced to practically nothing since all the forms of both the United States and the World Tests are printed in one twelve-page booklet and may be used over and over again.

A summary of the advantages of this type of test is as follows: (1) the places to be located are those which, in the judgment of experts, the average person ought to be acquainted with; (2) they are short and may be quickly given and scored; (3) they require a minimum amount of expressional effort on the part of the pupil, only one word or abbreviation being required in each exercise; (4) there are no debatable answers; (5) they are convenient for quick review to see where the pupils need help in place geography—whether cities, rivers, countries, etc.; (6) a large amount of testing may be done at a very small cost.

The author, of course, is aware of the fact that place geography is only one of the factors to be considered in measuring the results of geography teaching. The Bureau of Educational Research at Ohio State University has been making an intensive study of geography and has sought the advice of a large number of geography experts. The place-geography tests are therefore only a part of a proposed series of diagnostic tests.

Another part is a problem test which attempts to measure the ability of the pupil to apply geographical principles to concrete situations. Still another part is an information test[5] which is intended to measure the pupil's knowledge of important human relationships. A study is also being made to determine the technical vocabulary necessary for reading geography texts, geographical readers, and current news articles. When this study has been completed the important words will be embodied in a vocabulary test.

PRACTICE EXERCISES ON MINIMUM ESSENTIALS

Testing procedure has no real place in the public schools unless it helps in the improvement of instruction. Many enthusiastic workers have given tests and have found out where their pupils stand in relation to other pupils; but they have made no consistent effort to remedy matters. The author is preparing practice exercises in place geography to help the teacher emphasize the items upon which the pupils have failed.

These exercises are based upon outline maps of the United States and the World. The names of states, continents, bodies of water, etc., are not given, but upon the respective items numbers are placed. For example, the states in the United States map are numbered from 1 to 48 inclusive. Each pupil is given an outline map. With this map in hand the pupil may answer the teacher's questions either orally or in writing. The answer may be the number of the state as indicated on the map, the abbreviation of the state, or the name of the state. In rapid written exercises the number of the state has been found to be a great economy of time. Twenty-five to fifty questions may be answered by the pupils in five minutes. The papers may also be corrected in a very short time. The pupil corrects his own or his neighbor's paper as the teacher reads the correct answers. Some of the suggested practice questions on place and informational geography are as follows: (The answer may be the number, the abbreviation, or the name of the state according to directions given to the pupils.)

---

[5] A problem-information test on the United States will soon be ready for distribution. It will be published by the Public School Publishing Company, Bloomington, Illinois. Douglas C. Ridgley of Clarke University is cooperating with the authors in developing this series of tests.

1. What is the number of New York?
2. In what state is Chicago?
3. What state is south of Kansas?
4. Through what states would you pass in going directly from Kansas City to Denver?
5. Which state has the higher elevation, 1 or 20?
6. Which state mines the more coal, 25 or 13?
7. Which state has the larger percent of colored population, 16 or 48?
8. Which state has the warmer winters, 21 or 36?
9. In what state is Pikes Peak located?
10. In what state is the Great Salt Lake?

Other questions may be asked on relative distances or general directions such as:

1. What general direction is Alabama from Illinois?
2. About how far is Chicago from New York City? (In answering such questions a ruler may be used in connection with the scale given on the map.)
3. Which state is farther north, Oregon or New Jersey?
4. In what general direction does the Mississippi River flow?
5. In what part of the United States are the Rocky Mountains located?

Investigators who have studied methods of teaching spelling suggest that pupils be asked to write the words in a given lesson before studying them. Then, the words which a pupil already knows do not need to be studied. His time may be spent on the words which are new or difficult for him. The same time-saving method might very well be applied to place geography. The teacher, upon receiving pupils from a lower grade, should test their knowledge not only on the work already covered in the previous grade but also on the places to be studied under her instruction. Emphasis may then be placed upon the items with which the individual pupils are unfamiliar. Intelligently directed drill exercises of from five to ten minutes once or twice a week will doubtless be sufficient to fix these items in the pupils' minds.

In advocating the use of these practice exercises, the author does not consider it necessary that more time be spent on place geography, but he does suggest that the time now spent be used to better advantage.

THE INTELLIGENCE QUOTIENTS OF ADULTS AND RELATED PROBLEMS

WALTER F. DEARBORN
Graduate School of Education, Harvard University

With the exception of the results from the testing of intelligence in the United States Army, representative data have not been available from which it has been possible to estimate the average intelligence of age groups after the chronological ages of 13 or 14. From the ages of 7 or 8 up to 13 or 14 the age groups in schools are sufficiently representative. After these ages the groups are selected as a result of the elimination of pupils from school. In consequence of this fact mental ages over 14, as assigned in the Binet tests and in the various group intelligence tests are somewhat hypothetical. If, however, the age at which large numbers of children begin to leave school is really the age at which mental growth ceases *on the average,* as is suggested by the results of the Army testing, the solution of various problems of adult intelligence would be much simplified. This coincidence, if it be such, may, however, simply indicate that the tests in question measure chiefly the habits and interests ordinarily acquired in school, which many adults may either have never learned well or have forgotten. In their place other interests and habits may have been acquired, in other environments than the school, which are quite as significant for any adequate estimate of intelligence, but which the tests in question do not adequately test. That children of 12 and 13 in school do as well[1] on the average as the adults in the Army did in these tests may, for example, be due in part to the skill which children have acquired from recent practice in the use of paper and pencil, in the carrying out of directions, in writing, figuring, etc.,—a skill which many of the adults may have lost. On the other hand, the considerations presented by E. A. Lincoln in a recent article on "The Average Adult Mental Age"[2] seem to the writer sufficiently conclusive to place the average adult mental age, in so far as it can be determined by these tests, nearer to 13½ than to 16 as assumed by Terman in the Stanford Revision of the Binet Tests.

---

[1] See results of Proctor with the Army Alpha Examination in *Psychological tests and guidance of high school pupils.* (Journal of Educational Research Monograph No. 1); and DeVoss, *Manual of instructions for use with the Army Alpha Intelligence Examination in public schools,* p. 20.

The same limitations apply in good part also to the Binet tests. The latter are excellent measures of actual and potential grade progress. That they measure as successfully abilities expressed in other forms of learning is, however, doubtful. The limitations may apply somewhat less to the series of tests used in this study, because of the smaller proportion of the tests, for the ages in question, depending on scholastic training. Half of the tests and half the credits are for performances which may have been learned as well out of school as in school. Furthermore, the individuals included in the first part of this investigation were still in either the regular or continuation schools and may, therefore, have been in a somewhat more favorable position than were the adults in the Army tests. The determination of the proportion of those whose development has been inhibited by their school training, prolonged retardation in grades, etc., and for whom out of school learning would have been a fairer gauge of their intelligence awaits the development of more adequate tests than are now available.[3] It is generally recognized that mental tests do not measure the mental ability of an individual directly but gauge his ability by what he has learned or by the use to which he can put his learning in novel situations, and in either case a modicum of "drive" or of interest in the learning is assumed. His habits or acquisitions cannot, however, be appraised without taking into account the incentives and opportunities which he has had for their acquirement. In other words, environment cannot be left out of the reckoning; it is part of the equation.

---

[2] Lincoln, E. A. "The average adult mental age," *Journal of Educational Research*, 6:133-44, September, 1922.

[3] See in this connection a statement by the writer in regard to the measurement of intelligence in the *Journal of Educational Psychology*, April, 1921, p. 210.

The majority of tests in current use are reliable aids in the differentiation of the sort of abilities which the school fosters and prizes. On the basis of these results, alone, however, we should be chary of pronouncing judgments on the intelligence of those whose learning may have taken a quite different course.

The findings in regard to general physical and physiological development and such considerations as are presented by various writers[4] for placing the average age of full maturity at a higher age must be kept in mind. The writer, however, does not propose in this article to enter further into these discussions, but proposes, first, to present new findings in regard to unselected groups of fourteen- and fifteen-year-olds, and, secondly, to test the assumption of a lower mental age for the average adult by applying it to data secured from the testing of practically all of the children of three communities from the ages of 8 to 13 inclusive, and those who remained in school after these ages. A method will also be presented for calculating, on the basis of the assumed average, the intelligence quotients of individuals who score above this average. The values secured are directly comparable with those secured, in the case of scores below this average, by the usual method of dividing the mental by the chronological age. The method obviates the necessity for other indices or coefficients of brightness than the usual intelligence quotients.

One misconception, common at least to the lay mind, in speaking of the average mental age of adults may first be mentioned. If this age is assumed to be 14½, as will be assumed for the purposes of this article, it must be noticed, of course, that this is an average and that all of our findings would lead us to expect as great a range of variability on the one side of this average as on the other. Individuals of adult years are commonly known who, without the direct incidence of accident or disease, can only pass the mental tests of two-year-olds, one-year-olds, or less. At the other extreme we should expect to find individuals as much superior to the average as the one-year-olds are inferior to it. The final status of the individual variant and of the average is dependent on at least three factors, the initial status or native ability, the rate of growth, and the duration of growth. ("Environment" should also be added.) There is evidence of some positive correlation between these factors, that is, for example, that the development of initially superior minds is more rapid and distributed over a longer time. The recently published results of Kuhlmann[5] support this conclusion, although they indicate that the differences in the *duration* of the period of development are not as great as hitherto assumed.

---

[4] Otis, A. S. "A criticism of the Yerkes-Bridges point scale with alternative suggestions," *Journal of Educational Psychology*, 8:129-50, March, 1917; and "Absolute point scale for the group measurements of intelligence," *Journal of Educational Psychology*, 9:239-61, 333-48, May-June, 1918; Buckingham, B. R. and Monroe, W. S. *Illinois Examination. Teacher's handbook.* Bloomington, Illinois: Public School Publishing Company, 1920, p. 4; Brooks, F. D. *Changes in mental traits with age.* (Teachers' College, Columbia University Contributions to Education No. 116); Kuhlmann, F. "The results of repeated mental re-examinations of 639 feebleminded over a period of ten years," *Journal of Applied Psychology*, 5:195-224, September, 1921.

A more serious difficulty, in the application of the above assumption to our data, arises from the finding, which is common to the results of group testing, that the average increments in mental growth appear to be uniform up to this age; and it is difficult to maintain that it ceases abruptly. The standard scores for the ages 8 to 14 in Table I are almost the exact medians (with slight smoothing) of the age groups in question.[6]

TABLE I. AGE STANDARDS, DEARBORN GROUP TESTS OF INTELLIGENCE, SERIES II

Age	8½	9½	10½	11½	12½	13½	14½
Scores	25	37	49	61	73	85	97
Number of Cases	400	429	395	413	379	412	337

The average number of cases for ages 8 to 13 inclusive is 404, the number of fourteen-year-olds is 337. The decrease in the number of cases at 14 years is due chiefly to the elimination of children as they reach the age of 14, so that the fourteen-year-olds become a selected group. Since, as will be shown below, this elimination includes a greater proportion of those below than of those above the average in intelligence, those remaining are a superior group. The increment in score is due, therefore, in part to selection and in part to growth. In view of this fact it may be argued that the average rate of growth is decreasing at this period.

---

[5] Kuhlmann, F., *op. cit.*

[6] The scores are those which have been adopted as standards for the revised edition of Series II of the Dearborn Group Tests of Intelligence. The method of standardizing rested on other considerations than the exact medians and is described in the *Journal of Educational Psychology*, May, 1922. The median scores for the age groups in question are as follows: *8*, 25; *9*, 35; *10*, 48; *11*, 64; *12*, 72; *13*, 85; *14*, 99.

There are also other considerations to be taken into account. The average in question is a combined average for boys and girls. It is evident from some preliminary analysis that there are considerable sex differences at this period.[7] The average of the two curves may be considerably different from that of either alone. Moreover, it is possible that there are abrupt changes at this period in the individual growth curves. This can be proved only by repeated tests of the same individuals. An increase observed in the variability of the age groups at ages 12, 13, and 14 is presumably due to this factor and to the sex differences. The changes in variability are shown in Table II in terms of the semi-interquartile ranges of the scores, mental ages, and intelligence quotients.[8]

Finally, the form of the curves may be varied at will depending on the nature and difficulty of the tests given at various ages,

---

[7] See chapter on sex differences in L. Thomas Hopkins' *The intelligence of continuation school pupils in Massachusetts* to be published in the Harvard Studies in Education. A more extended investigation of sex differences in these tests is now being made by Mr. E. A. Lincoln.

[8] It should be noted in passing that the commonly accepted averages for other forms of growth, e. g., physical and physiological development, are probably in need of some correction after age 13, since they have also been based only on children who remain in school. How much of the increment in height after age 13 is, for example, due to the factor of selection? Presumably not much, but this possibility must be considered by those who argue in regard to mental growth on the analogy of the usually accepted averages in regard to physical growth. See, for example, the effects of selection in the differences between the "Längenwachstum von Berliner Gymnasiasten und Gemeindschülern," W. Camerer in *Handbuch der Kinderheilkunde,* p. 244. Also Baldwin, B. T., *Physical growth of children from birth to maturity,* (University of Iowa Studies, Vol. I, No. 1) Table III, p. 32. Fourteen-year-old children granted working certificates were at least equal to the norms, the fifteen-year-olds are under the norms in height and weight.

TABLE II. VARIABILITY OF SCORES, INTELLIGENCE QUOTIENTS, AND MENTAL AGES

Age	8	9	10	11	12	13	14	15	16
Score	10.2	11.5	14.8	15.2	20.3	22	20.1	17.7	15.
I. Q.'s	10.1	10.	11.7	11.4	13.	13.7	11.6	9.8	9.2
M.A.(months)	11.0	11.5	14.5	14.5	20.5	22.	20.	17.5	15.5

and the points of score considered as units. This would also be true of the Binet but for the fixed method of counting units. A year of mental growth in the Binet tests is counted as one unit, although the amounts of such growth in one year may vary at different periods. The only information required in the following discussion is at what age period there is on the average no advance over other age periods. This last consideration may be illustrated by the graphic comparison (Figure 1) of the *scores*

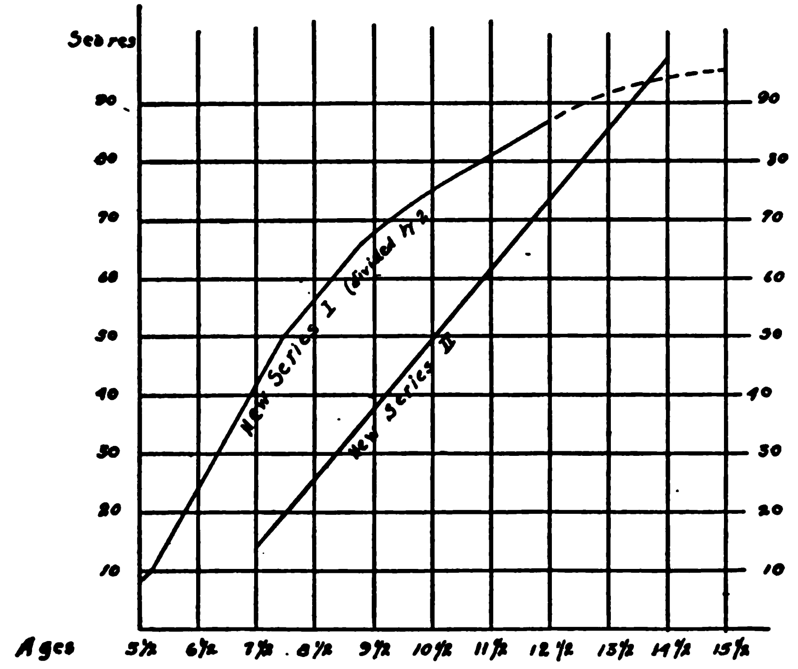


FIGURE 1. AGE STANDARDS, DEARBORN GROUP TESTS OF INTELLIGENCE, NEW SERIES I AND II

in Series I and II of the writer's group tests. Reliable averages for the First Series have been obtained only up to the age of 12; the dotted line indicates the probable course of the curve after this age. The scores of this Series have been divided by two in order to make the comparison easier with the scores of Series II.

To turn now to the findings on the first mentioned data. The standard scores and mental ages in both individual and group tests in current use are based almost exclusively on the results obtained from children in the regular grades and high schools. The Massachusetts laws require that children who leave the regular schools between the ages of 14 and 16 years must attend continuation schools. If we are to secure an unselected group on which to base our average, these children must be included with those in the regular schools. An investigation now being made at the Harvard Graduate School of Education by Mr. L. T. Hopkins[9] has supplied this necessary information in regard to practically all of the fourteen- and fifteen-year-olds in several Massachusetts cities. In one typical city, there are found to be 454 fourteen-year-olds and 457 fifteen-year-olds. The mean scores of the fifteen-year-olds in the Dearborn Group Tests[10] show no appreciable advance over the fourteen-year-olds. The scores and mental age equivalents are shown in the last two columns of Table III.

If one considers those in the regular schools alone as has hitherto been done, the fifteen-year-olds advance over the four-

TABLE III. MEAN SCORES AND MENTAL AGES OF PUPILS FOURTEEN AND FIFTEEN YEARS OLD

	IN REGULAR SCHOOL		IN CONTINUATION SCHOOL		IN ALL SCHOOLS	
	(1) Age 14	*(2)* Age 15	*(3)* Age 14	*(4)* Age 15	*(5)* Age 14	*(6)* Age 15
Scores	119	129	81	89	110	111
Mental Ages	14–5	15–1	11–10	12–4	13–9	13–10
Number of Pupils	342	249	112	208	454	457

---

[9] *The intelligence of continuation school pupils in Massachusetts.*

[10] The scores are of Series II, General Examinations 4 and 5. Scores and comparisons made elsewhere in this article are, except as noted, based on the Revised Series II, General Examinations C and D.

teen-year-olds (columns 2 and 1 of Table III). This is the resultant, as pointed out, of the elimination of the continuation-school group. There are 112 of the fourteen-year-olds and 208 of the fifteen-year-olds in the continuation school. Their average scores are respectively 38 and 40 points less than those in the regular schools. In terms of mental age they average, respectively, 2 years and 7 months, and 2 years and 9 months lower than the pupils in the regular schools.

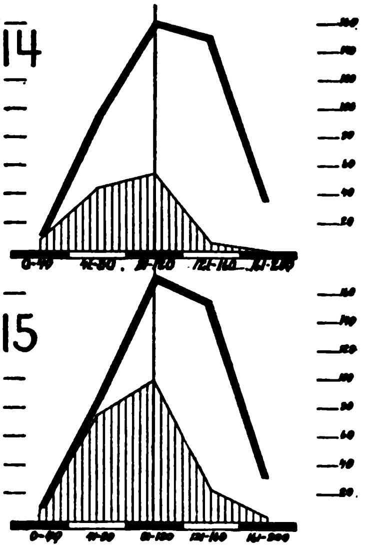


FIGURE 2. SCORE DISTRIBUTIONS BY AGE IN THREE MASSACHUSETTS TOWNS. SERIES II, GENERAL EXAMINATION 4 AND 5

In Figure 2 the distributions of the scores of all the fourteen-year-olds and of all the fifteen-year-olds are shown by the heavy lines. The shaded areas show the scores of the continuation-school groups. The similarity in the form of the two distributions is striking. The only children of these ages in the cities who are not included are those in the private and parochial schools. The records returned by these schools show that there are 126 of these children. The fourteen- and fifteen-year-olds are not separated in these returns; but they must be approximately equal, judging from equality in the numbers of these ages in the public schools. Possibly these children might somewhat raise or lower the average scores of both groups; but probably they would not affect the similarity of the groups nor the equality of their averages because, aside from the small number of cases involved, it is not likely that the sort of fourteen-year-old children going to private and parochial schools differs materially from the sort of fifteen-year-olds going to these schools.

The average mental ages of these two total groups in terms of the standards of the Dearborn Group Test are 13 years 9 months and 13 years 10 months respectively. The average mental ages of some 3400 fourteen- and fifteen-year-olds in the cities so far studied are practically the same; namely, between 13 years 8 months and 13 years 10 months. If these cities were typical of the country at large, we should need to assume that the average adult age, as far as these tests show it, is between thirteen and a half and fourteen. For various reasons, chiefly because of the proportion of children of foreign born parents, we do not believe these communities, or at least the results of the tests, are quite typical of the average population of the country—hard as that is to estimate. The discussion in connection with Table I also pointed to an average age nearer fourteen, but for the purposes of the present argument, we shall, as noted above, assume the average adult age to be somewhat higher; namely, fourteen and a half.

If we make this assumption and change our scores (by a method described in the appendix to this article) into intelligence quotients, we have two ways of calculating the quotients. We may calculate them in the usual way, by dividing the mental age equivalents of the scores by the chronological ages of the pupils, or the fourteen- and fifteen-year-old scores or score equivalents may be divided by no chronological age higher than fourteen and a half.

The results in the case of the fourteen-year-olds are (omitting decimal points) practically the same by either method. But if we use the usual method for the fifteen-year-olds, those in the regular schools have a slightly *lower* median I.Q. (101 instead of 102) than the fourteen-year-olds, although we know that, because of the removal of pupils to the continuation school, they must average higher than the fourteen-year-olds. The total group of fifteen-year-olds average but 90 I.Q. as compared with 96 for the fourteen-year-olds, whereas they should average the same. Calculating the I.Q.'s on the basis of 14½ as the average adult age, the findings are such as the facts established by the scores would indicate; namely, first a gain in the average I.Q.'s of the fifteen-year-olds in the regular schools of five points, and secondly, about the same average I.Q.'s, 96 and 95, for the total groups of fourteen- and fifteen-year-olds. The different results by the two methods are shown in Table IV and in Figure 3. The broken lines in the figure show the results by the usual method, the continuous lines by the proposed method. The I.Q.'s are plotted in intervals of 15 points.

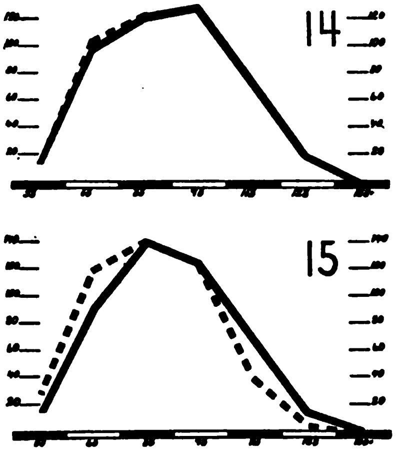


FIGURE 3. I.Q. DISTRIBUTIONS FOR FOURTEEN- AND FIFTEEN-YEAR-OLD PUPILS IN ONE MASSACHUSETTS TOWN. SERIES II, GENERAL EXAMINATION 4 AND 5. SOLID LINES ARE THE RESULT OF TAKING THE ADULT AGE AS 14 YEARS AND 6 MONTHS. THE BROKEN LINES ARE THE RESULT OF TAKING THE ADULT AGE AS 16 YEARS

TABLE IV. MEDIAN INTELLIGENCE QUOTIENTS OF FOURTEEN- AND FIFTEEN-YEAR-OLD PUPILS

	In Regular Schools		In Continuation Schools		In All Schools	
	Median	Quartile Deviation	Median	Quartile Deviation	Median	Quartile Deviation
Fourteen.........	102	12	83	16	96	13
Fifteen						
(a) Average Adult Age = 16........	101	10	80	7	90	12
(b) Average Adult Age = 14½.......	106	10	85	8	95	13

(Median Score= 85. ; —Q=22. ; +Q=22.
Median I.Q. =100.5; —Q=13.8; +Q=13.8)

The method will now be applied to the age and grade distributions of three communities in which the tests were given from the second grade through the high school. A study of the effect in these communities of changes in population and of school elimination and retardation warrants the conclusions that the pupils who remain in school after the age of 14, and after the seventh grade, are considerably superior to the preceding age and grade groups. The calculation of the I.Q.'s of these groups on a basis of 14½ as the average adult age, gives results in closer agreement with this generally accepted conclusion, than by calculating on the basis of sixteen years or over.

The chief findings in regard to the influence of elimination and retardation for the single year investigated are as follows:

1. The median intelligence quotients of those who left school in the three towns was 100 for ages 10 to 13 inclusive (81 cases), whereas it was 90 for ages 14 to 16 inclusive (109 cases).

2. The median intelligence quotient of those given work certificates in two of the towns at the age of 14 and over was 89 (21 cases). The facts for the third town were not available.

3. The median intelligence quotient of those not going on to high school in this third town was 84 (17 cases).

4. The median intelligence quotient of those failing of promotion in one town in the year since the tests were given was 87. The median has not been established for the other towns, but it is believed to be considerably lower.

It has not been possible to study thoroughly the effect of possible changes in the population, but such inquiries and observations as could be made seem to support the conclusion that, taking the three towns together, there has not been any marked change in the character of the population and that, therefore, the differences between the average intelligence of pupils in grades ii to vii, and viii to xii, and before and after age 14, are due to the selective influences of elimination and retardation.

These differences are shown in the following Table V:

TABLE V. MEDIAN I.Q.'S AND VARIABILITIES BY AGE AND GRADE IN THREE MASSACHUSETTS TOWNS

Grade	Number of Cases	Basis of Calculation				Age	Number of Cases	Basis of Calculation			
		Usual		Proposed				Usual		Proposed	
		Median	Q	Median	Q			Median	Q	Median	Q
II	380	97.9	8			7	235	101.7	7		
III	401	96.4	10			8	388	101.4	10		
IV	409	97.6	11			9	417	99.0	10		
V	407	100.8	10			10	377	100.5	12		
VI	387	101.3	11	101.3	11	11	387	102.9	11		
VII	416	99.3	13	99.4	13	12	385	100.2	13		
VIII	325	104.2	12	104.8	11	13	397	100.5	14		
						14	332	101.5	11	101.8	12
IX	286	106.5	10	108.1	9	15	257	101.6	9	107.6	10
X	265	105.6	8	110.7	7	16	223	102.0	8	111.7	9
XI	203	107.0	7	114.5	8	17	101	105.1	8	115.0	8
XII	131	108.1	7	119.7	8	18	37	103.8		114.3	

The average I.Q.'s of age groups up to 13 inclusive and of grades v to vii inclusive are approximately 100. The lower medians in grades ii, iii, and iv are due to the retarded pupils. On the usual basis of calculation, the I.Q.'s of the fifteen-year-olds average the same as the fourteen-year-olds, namely about 101. On the proposed basis the I.Q.'s of the fifteen-year-olds average 107, the sixteen-year-olds nearly 112, and the seventeen-year-olds 115. The median I.Q.'s of the high-school pupils advance, by the usual method of calculation, but little over the I.Q. of the eighth grade; 104 for the eighth grade and 108 for high-school seniors. By the proposed method the high-school seniors average about 15 points higher in their intelligence quotients than the pupils of the eighth grade. These comparisons are best shown by the accompanying Figures 4 and 5 in which the dotted lines

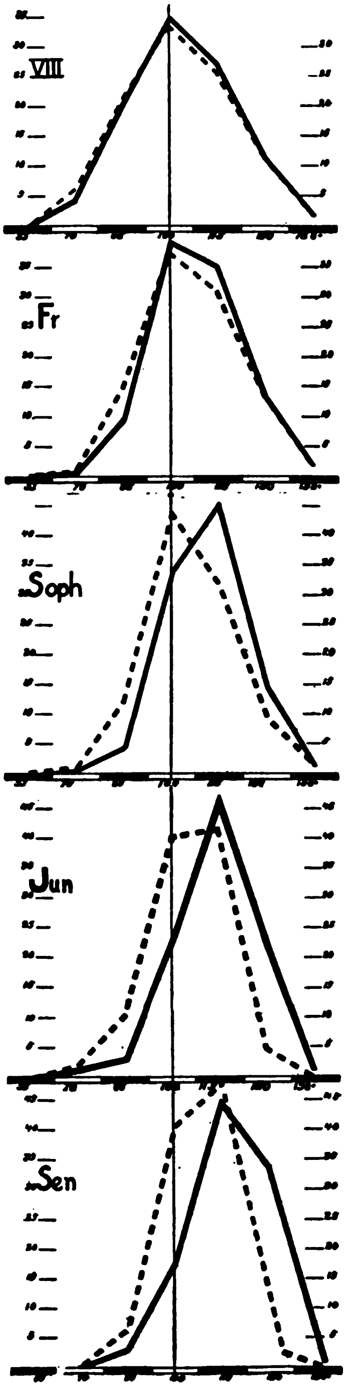


FIGURE 4. I.Q. DISTRIBUTIONS BY AGE IN THREE MASSACHUSETTS TOWNS. SERIES II REVISED. FULL - LINE CURVES RESULT FROM TAKING ADULT AGE AS 14 YEARS AND 6 MONTHS. BROKEN LINES SHOW RESULTS FROM TAKING ADULT AGE AS 16 YEARS.

show the distributions of I.Q.'s by the usual method and the continuous lines by the proposed method.

The findings of Proctor with the Army Alpha, by presumably the usual basis of calculation, are of interest by way of comparison. He finds the median I.Q. of the first year high to be

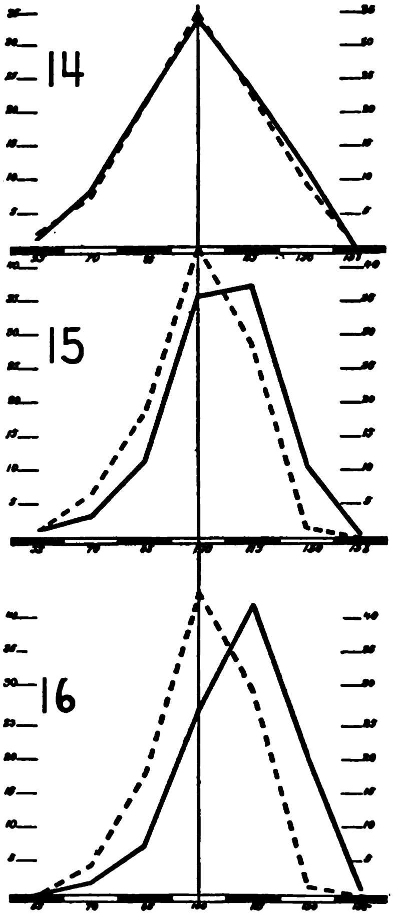


FIGURE 5. I.Q. DISTRIBUTION BY GRADES IN THREE MASSACHUSETTS TOWNS. SERIES II REVISED. FULL-LINE CURVES RESULT FROM TAKING ADULT AGE AS 14 YEARS AND 6 MONTHS. BROKEN LINES SHOW RESULTS FROM TAKING ADULT AGE AS 16 YEARS

105, of high-school graduates 111, and of those going to college 116,[11] and adds that "if the same process of selection takes place in college that we have shown in high school we should expect the median I.Q. of college graduates to be 120 or over."[12]

The tests used in this study[13] have also been applied to a number of undergraduates and graduate classes at Harvard University with the following results calculated on the proposed basis:

Class composed chiefly of freshmen	(54 cases)	I.Q. 125
Classes of upper-class men	(80 cases)	I.Q. 133
Classes of graduate students	(120 cases)	I.Q. 137

CONCLUSIONS

Evidence has been presented which shows that the average adult mental age is approximately 14½ years and that the calculation of the intelligence quotients on this basis gives results in closer agreement with established facts than by the method in current use. The conclusions of this study are believed to be more reliable than those based on the Army Alpha and Binet tests, because the tests used in the study are better suited to the purpose, the main groups studied more representative, and the conditions under which the tests were taken fairer for the comparison than in the case of the Army testing. A method is presented in the appendix for the calculation of the intelligence quotients of individuals who are superior to the average adult.

APPENDIX

Method of Translating Adult Test Scores into Intelligence Quotients

The intelligence quotients, which are obtained by the following method, express the relative brightness of individuals who are superior to the average adult in the same way and have the same value as those obtained below this level by dividing the mental age by the chronological age of the individual. The method has the advantage of using the same index of brightness for adults as for children, and will be particularly useful in studying such problems as the constancy of intelligence quotients, where it is desirable to compare the relative brightness of individuals during the period of growth and after maturity.

---

[11] *Op. cit.*, p. 22.

[12] *Op. cit.*, p. 59.

[13] In these classes old Series II (General Examinations 4 and 5) was used.

For purposes of illustration the method is applied to the scores of the thirteen- and fourteen-year-olds of the three communities which are described in the body of this article and which are believed to be fairly "average" of typical communities. The thirteen-year group includes all children in these towns between the ages of 13-0 and 13-11 inclusive with the exception of a very small number in private schools and a negligible number of children of very low mentality. The latter group is somewhat selected because of the elimination from school of a slightly larger proportion of pupils of inferior intelligence than of superior intelligence. If the tests used are valid measures, the average of the latter groups has reached maturity.

The median score of the thirteen-year-olds in the revised Series II of the writer's group tests of intelligence is 85. A pupil securing this score is considered, following the usual practice, to have a mental age of thirteen and a half years. The median score of the fourteen-year-olds is 98.8. The score established as standard for mental age fourteen and one-half years is a little lower, namely 97, because, as just noted, this group is somewhat selected. The mental age standards up to and including age 14 have thus been established by the median accomplishments in the tests of representative groups of the various chronological ages. It is not possible, however, by this latter method to establish mental ages beyond 14, because, as argued in the body of this report, the median scores of unselected groups of ages 15 and above do not seem to advance above the median of the fourteen-year-olds.

It is possible, however, to fix upon points in the distribution of scores above the fourteen-year median which have the same value and significance as the mental ages below this median. These points or scores may, for convenience, be designated as mental ages 15, 16, etc., and the latter may be used to calculate the intelligence quotients of individuals who are superior in intelligence to the average adult or median fourteen-year intelligence.

In the following presentation of the method these intelligence quotients are determined directly from the scores in the tests. It is equally feasible to obtain first the mental age equivalents to the scores and then figure the intelligence quotients from them.

The distributions of the scores of both the thirteen- and fourteen-year-olds are approximately symmetrical, as is indicated in Table VI. The difference between the median and the 25-percentile is, in the case of the scores of the thirteen-year-olds, 21.95. This difference or distance along the base line of the distribution is designated in the following discussion as "— Q". The difference between the median and the 75-percentile is 22.02. This difference is designated as "+ Q." In the case of the 14-year-olds "— Q" = 20.85 and "+ Q" = 19.40.

TABLE VI

	13 Year Distributions		14 Year Distributions	
	Score	I. Q.	Score	I. Q.
Median	85.10	100.5	98.81	101.8
25-Percentiles	63.15	86.7	77.96	90.0
75-Percentiles	107.12	(?)	118.21	(?)
—Q	21.95	13.8	20.85	11.8
+Q	22.02	(?)	19.40	(?)

The problem, as indicated by the question marks in Table VI, is to determine the "+ Q" distances and the 75-percentiles of the I.Q. distributions.

The intelligence quotients of all the thirteen-year-olds who secured scores equivalent to mental ages up to and including 14½ were first calculated by using the standards given in Table I of the foregoing article. Since the intelligence quotients of individuals having scores over that equivalent to mental age 14½ are all well over 100 (the lowest is above 103.6), they are placed above this point in the frequency distribution and the median I.Q. and the 25-percentile of the I.Q. found. The median I.Q., as given in the above table, is 100.5; the 25-percentile is 86.7; "— Q" therefore equals 13.8.

It can now be shown that for a chronological age of 13½ a score below the median can be directly translated into an intelligence quotient by taking the same "Q" location on the two distributions and that the result is the same as it secured when the intelligence quotients are calculated by reference to the standards of Table I. (It may be noted again that the standards below age 13 are based on the median scores of different groups of pupils than the one now under discussion, namely, children of ages 7 to 12 inclusive.) Since there is this correspondence *below* the medians of the two distributions, and, since the distribution of *scores* above the median has practically the same variability as that below, it is fair to assume that the intelligence quotients above the median would also correspond in the same ways to the scores.

The following examples will illustrate the above statements:

A score of 63 in these tests is equivalent, according to the standards of Table I, to a mental age of 11 years, 8 months, and, for a thirteen and a half-year-old, to an I.Q. of 86. This score, as may be seen from Table VI, is at the 25-percentile point of the distribution of thirteen-year-old scores. The corresponding (25-percentile) point of the I.Q. distribution is 86.

The 75-percentile score is 107. Since this score is by the percentile count as much above the median as 63 is below the median, it should have for its equivalent an I.Q. as much above the median I.Q. as 86 is below it, namely 14 points, or 114. The 75-percentile point of the I.Q.

distribution may be secured by adding the amount of "—Q" to the ***median*** of the thirteen-year intelligence quotients. This is done on the assumption that, since the distribution of *scores* is approximately symmetrical and "—Q" of the scores practically equals "+ Q" (scores), the upper half of the intelligence quotients is also similarly distributed, and that, therefore, "+ Q" of the I.Q. distribution (which is unknown) would equal "—Q" of this distribution (which is known). "—Q" of the thirteen-year I.Q. distribution is 13.8 which, added to the median (100.5), gives 114.3. This, to repeat, is, as was required, as much above the median in I.Q. points as 86 is below it.

When, as in the fourteen-year-old distribution of scores, "+ Q" differs from "— Q" (the distribution of scores not being quite as symmetrical as in the above case), close approximation to the measure of the variability of the intelligence quotients above the median, or "+ Q", may be secured by taking the amount which $\frac{+Q}{-Q}$ of the distribution of scores is of "— Q" of the distribution of intelligence quotients. In the case of the fourteen-year-olds (See Table VIII) this is as follows:

$$\frac{19.40}{20.85} \times 11.8 = 10.97 = \text{"}+Q\text{"}.$$

This correction has been made use of in Table VIII. In Table VII, "— Q" and "+ Q" are taken to be equal.

We may now equate all the scores and I.Q.'s of the lower half of each distribution by successively subtracting these quartile ranges (or fractions thereof) from their respective medians until the lower limits of the distributions are reached. The *points* on these distributions thus found are for convenience designated as — 1Q, — 2Q, — 3Q, — 3.5Q. As a check on the method, the intelligence quotients which have been calculated from the standard mental age equivalents of the scores are also given. This comparison is made in the upper half of Table VII. As will be seen the results are practically the same by either method. The scores and I.Q.'s corresponding to the intermediate points between — 1Q and — 2Q, between — 2Q and — 3Q, etc., may be equated by taking equal fractions of the respective Q distances.

The I.Q. equivalents of the scores above the median may be found by successively adding the quartile ranges or fractions therof to the respective medians. These are designated in the Table, + 1Q, + 2Q, + 3Q, + 3.45Q. The last point equals the maximum score on the tests of 171. The figures are given in the lower half of Table VII. The maximum score of 171, an I.Q. of 148, and for age 13.5 a "mental age" of 20, are as much superior to the median performance for this age as a score of 8, an I.Q. of 52, and a mental age of 7 are inferior to this median.

TABLE VII. I.Q. EQUIVALENTS OF TEST SCORES OF AGE 13½

	By Q. Method		By Age-Score Standards		
	Scores	I. Q.	Scores	M. A.	I. Q.
−1Q	63	86.7	63	11–8	86
−2Q	41	72.9	41	9–10	73
−3Q	19	59.1	19	7–11	59
−3.5Q	8	52.2	8	7–0	52
+1Q	117	114.3			
+2Q	139	128.1			
+3Q	161	141.9			
+3.45Q	171	148.1			

(Median Score= 85. ;—Q=22. ; +Q=22.
Median I.Q. =100.5;—Q=13.8; +Q=13.8)

The same methods and comparisons for age 14 are shown in Table VIII. The "+ Q" of the I.Q. distribution has been taken according to the above calculation as equal to 10.9. The I.Q.'s according to the age-score standards are about one point lower than by the Q method. As noted the fourteen-year-olds studied are a somewhat selected group, the check by method B would be exact if the median I.Q. were taken at 100.8 instead of 101.8.

TABLE VIII. I.Q. EQUIVALENTS OF TEST SCORES OF AGE 14½

	By Q. Method		By Age-Score Standards		
	Scores	I. Q.	Scores	M. A.	I. Q.
−1Q	78	90	78	12–11	89
−2Q	57	78.2	57	11–2	77
−3Q	36	66.4	36	9–5	65
−3.5Q	27	60.5	27	8–8	60
+1Q	118	112.7			
+2Q	138	123.6			
+3Q	157	134.5			
+3.5Q	167	139.9			
+3.72Q	171	142.3			

(Median Score= 98.8; —Q=20.8; +Q=19.4
Median I.Q. =101.8; —Q=11.8; +Q=10.9)

ADJUSTMENTS IN CALIFORNIA[1]

Will C. Wood
State Superintendent of Public Instruction, Sacramento, Californi

In appearing before you this afternoon I feel like one deals in faith appearing before a group of diagnosticians.

As Superintendent of Public Instruction during the last years it has been my business to renew people's faith in educa I must congratulate the State of California on the numb people in this meeting today. Something has happened in fornia in recent years. Ten years ago if we had gotten tw people together at a voluntary conference like this we would considered ourselves very fortunate indeed. There has bee educational awakening and I believe it is the most hopeful we have in California—the fact that people recognize the something beyond what we now know that we want to about.

Our educational program has been sold in California successfully. The people of California have shown by a mendous majority that they desire to keep the schools on a p basis. As I go through the state addressing various grou business men, I find them greatly interested in education; still believe what we sold them is worth while. Howeve time is coming when we shall have to prove we are "deliv the goods" and if we cannot prove it, we shall find that th vance we think we have made has not been a real advance all. How are we going to prove that we can "deliver the go or that we are going to "deliver the goods"? It is important we measure the goods that are being offered in California t If we can measure the goods and prove to the people we ar livering full measure, then, indeed, we can hold to our e tional program and our salesmanship will have justified i

Last year ten percent of the boys and girls enrolled in th mentary schools of California were repeaters. Over seventy lion dollars was spent on education in California last year. Ten percent of this seventy million dollars was spent on re-education of repeaters. I am convinced we shall have to find some way of cutting down the number of repeaters when the tax payers begin to examine our work more carefully. Is it necessary to have ten percent of our young people each year repeat their work? Our tests should be able to throw some light on the problem of repeaters. I am inclined to think two percent of the repeating may be due to feeblemindedness; I am wondering about the remaining eight percent. Why are they repeaters? Somebody engaged in educational measurements will have to tell us why.

---

[1] Address delivered May 30, 1922, before the Conference on Educa Research and Guidance at San Jose, California.

Tax payers, when they get around to it, will also inquire why we have many classes of only twenty-five pupils in our California schools. Why shouldn't we have forty pupils? Some cities have fifty; others have only twenty-five. Why the difference? We shall have to be able to prove it is economical to have classes of twenty-five or the tax payers will draw conclusions that there is waste. How many pupils can a teacher handle successfully? Again educational tests ought to be able to throw some light on this problem.

We are paying salaries in California that justify a fairly long school term. Should the school term be three hundred, or two hundred, or a hundred and sixty days? Someone should be able to tell us how long the school term should be. Then again, how long should the school day be? Should it be six hours or five hours or three hours? Educational tests when properly developed should be able to throw some light on such administrative problems.

Certainly there is an ideal arrangement concerning these matters I have mentioned. Who will define the ideal? The way is open to any enthusiastic student of education to devise tests that will help answer these questions.

We hear of the efficiency or inefficiency of one-room country schools; also of the efficiency of consolidated schools. Do consolidated schools show better results than the one-room school? Tests ought to throw light on this problem. If we can prove

better results in the consolidated school, the problem of con
dation will be simplified. Just now when we go out to se
consolidation program to a community we have to deal too n
in theory and promise.

Should we have supervision in the rural-school districts?
schools have supervision. Some say there is too much. Are
wasting money on supervision in cities? Does supervision p
If so, how much? Some of these educational testers should
out and let us know. If we are going to have rural-school su
vision and retain city supervision it will be necessary for u
prove results.

Shall we teach penmanship one-half hour a day for three
a week or shall we teach it fifteen minutes a day for five da
week? I wish you would let me know. After one has atta
a certain degree of efficiency in penmanship isn't he a g
enough penman to be let alone? Doesn't the law of diminisl
returns apply in penmanship just as it does in agriculture? A
a pupil has attained a mark of seventy or eighty on Ayres' s
shouldn't we let him stop practicing writing and put the tim
on something else? We want our educational testers to tel
something about this problem.

Can you prove that we shouldn't have grammar in
seventh and eighth grades and thus meet the criticism of h
school teachers who are demanding it for admission to l
school? Can you not throw light on the problem when the st
of formal grammar should begin? Can you prove that t
spent in oral language is well spent?

Then there is the question of county examinations. I tl
if we had a bureau of educational tests properly organized
could prove that county examinations are not worth while
are not real tests of merit. Until we can prove their lacl
worth, until we can find something that will demonstrate
these examinations do not really test, school people will cont
to have faith in them.

Educational testers will also have to help us in the adminis
tion of the Teacher Tenure Law. A permanent teacher ca

be dismissed in California except for immoral or unprofessional conduct, incompetence, evident unfitness for teaching, or persistent violation of or refusal to obey the school laws. How are you going to prove incompetency without some kind of a test for teachers? My own belief is that the Teacher Tenure Law will be repealed at the behest of public opinion unless someone finds a means of testing the competency of teachers.

I wish someone would devise a scheme for measuring the ability of music teachers. They come to us asking for certification with all sorts of training—usually taken under some individual from Russia or England or some other foreign land. We have never heard of the instructor under whom the applicant trained. When we attempt to rate the qualifications and find them insufficient, the candidate invariably says it is a case of a layman passing judgment on an artist. I do wish someone would provide us with a means for testing abilities of artists. We should find someone who can tell us what is the value of music in education; also what is the value of the training that teachers of music are given.

We have a law providing that physical education shall be given a certain amount of time each week. We have had it since 1917. Is it worth while? Someone should be able to tell us whether results in physical education justify the expenditure for that purpose. I told the physical education people at their last meeting if we are to continue giving the time required by law to physical education, they must find some way of proving that the time is valuable in results. We are spending over $2,000,000 a year for physical education in California. Are we getting results? We can't take physical education on faith all the time. Someone said the other day that university students are stronger since we adopted our physical education program; but how much stronger? We ask the testers to tell us.

Now as to textbooks, isn't there some way of measuring the desirability of textbooks for elementary schools? I believe it can be done. We have to measure them now but our measurements are necessarily crude. We find many who say the geography we have adopted is a splendid book. We find others who say it is an atrocious text. Which opinion is right? It is time for us to find some way of measuring the value of texts.

We need tests of buildings and equipment as Professor Hart indicated here this morning. Certainly you would agree with me in this statement if you were to visit some of the rural schools I have visited recently. The county superintendents have no satisfactory standard with which to measure buildings and grounds. We ought to be give them such a standard. Professor Williams has told us of his plan for measuring homes. Won't someone make a standard for measuring school buildings and grounds, and schools as social institutions?

I have mentioned a few things that should be done in the line of measurements in California. I hope there will be other conferences of this kind so I have told you of a few things we would like to have you study. We have made a good beginning in California. We have our city bureaus and one or two counties are engaged in something in the way of measurements; we also have our university bureaus at Stanford and the University of California; we have our teachers' college research centers. Somebody said to me the other day, "What we need is a great central state bureau that will send out people to do the measuring." I do not think so. I do not believe in building up a great bureaucracy at Sacramento. I believe in stimulating local centers to do for themselves what they ought to do. Instead of having dozens of supervisors going out from Sacramento, covering the ground once in three or four years, I would have supervisors and research workers in counties and cities and teacher-training institutions. If the state tries to do research and supervision on a large scale and covers the ground, it will keep counties and cities out of the field. Because the work is so gigantic, the state would do it very poorly if it attempted it. I should like to see every county employ an agent whose business it would be to conduct educational measurements. However, we should need to have a state leader whose business it would be to serve the places which are not served by local people—someone to sell the idea of measurements to county superintendents, to gather results of observations and measurements, and to provide a clearing house where data gathered throughout the state might be presented to the advantage of others who are working on the same problem. The Sacramento center in the state office would not be a big center; it would be a stimulating center—a directing center. Until we can do this we shall not be able fully to carry out the great program we have undertaken in California.

I am glad to be with you and I am sorry I cannot tell you much about adjustments. I must content myself by telling you what some of the problems are. I think the work in which you are engaged is a wonderful work. I expect it to serve education by showing the people some of the things in the school system that are wrong and by assisting us to remedy them. I also expect your work to serve the school system by indicating to the people that the work we are doing in other respects is worth while; that the amount of money they are putting into education is bringing results commensurate with the amount they are spending. This, I believe, is one of the best ways in which educational measurement may serve the schools of our state.

A COMPARISON OF THE ACHIEVEMENT OF HIGH SCHOOL AND UNIVERSITY STUDENTS IN CERTAIN TASKS IN CHEMISTRY

S. R. Powers
University of Minnesota

Near the close of the school year of 1919-1920, a test of ability to perform the more or less mechanical operations of chemistry was given to a large number of high-school and university students. Twenty high schools participated, fourteen being in medium or small Minnesota cities and the remaining six being outside of Minnesota. Two groups of university freshmen were represented: (1) those who had taken one year of chemistry in the university, but who had not taken chemistry in high school; and (2) those who had taken one year of chemistry in the University and who, in addition, had offered chemistry as an entrance subject. The first group consisting of 57 students is designated in the tables below as U_1, the second group (168 students) is designated as U_2[1]. Table I shows the number of schools which used the test and the number of students who took it.

TABLE I. NUMBER OF SCHOOLS USING THE TEST AND NUMBER OF STUDENTS WHO TOOK IT

	Schools	Students
High Schools.................	20	731
U_1..........................	1	57
U_2..........................	1	168

The nature of the test is made clear from the following illustrations. These were the directions for each of the eight divisions just as they were printed on the test. Each division consisted of either five, ten, or twenty different items. All the items in a given division were of the same general nature but for the sake of brevity only one item from each is included in the illustrations. The complete test consisted of 95 items.

---

[1] The U_2 group had had qualitative analysis, the U_1 group had had only one year of general chemistry.

Illustrations:

1. Write the names of the following substances:
   HgO. (Ten compounds constitute this division.)
2. Write the formula for the following substances:
   Potassium carbonate. (Ten substances constitute this division.)
3. The valence of the following elements and radicals are as indicated. Write the formula for the oxide, bromide, sulfate, and phosphate of each metal.
   Mg^{++}, O^{--} (The valence of five positive and four negative radicals are given and space left for 20 formulas.)
4. Arrange the following elements in order of the activity series: (Ten positive elements are given and 10 blanks provided for arranging them in series.)
5. Place a letter E before each substance listed here if it is an element or a letter M if it is a mixture or a letter C if it is a compound. (The names of 20 substances are given.)
6. Fill in the right-hand side and balance each of the following equations: HgO (heated). (Ten chemical changes are given.)
7. Write the equation for each of the following changes:
   The preparation of oxygen from potassium chlorate. (Ten chemical changes are given.)
8. Write one equation to illustrate each of the following types of reactions: Simple addition. (Five types of equations are asked for.)

The results from this testing furnish data which are valuable in at least four particulars. (1) They show very definitely the difference between the ability of high-school and university students to do certain specific tasks. Both the U_1 group and the high-school group had taken chemistry for one year. These data are, therefore, significant criteria for judging the relative efficacy of the instruction in chemistry in high schools and in freshman classes. (2) It is possible to compare the effectiveness of instruction given in small and large high schools. (3) The high-school data are important because the test was taken by such a large number of high-school students that the scores on each of the items may be considered fairly reliable standards for high schools of the types tested. (4) A basis is afforded for judging the adaptability of the subject-matter to the ability of high-school students. If students generally fail on tasks which have been included in their instruction we may well question whether those tasks are adapted to the ability of the students. When it is found that as many as 75 percent of high-school students are unable to do a certain task, one of two conclusions are inevitable: either (a) the instruction on that task has been poorly given, or (b) the task is too difficult. If it should be found that both the high-school and university groups are unable to do a task upon which they have received instruction, this fact would furnish an additional basis for questioning the appropriateness of the given task.

Comparison of University Freshmen and High-School Students

The data in the tables below show the amount of overlapping in different schools. Table II shows the median and quartile scores,[2] and the range of scores for all high schools taken collectively. It gives the scores for those schools which ranked, on the basis of medians, as first (highest), second, fifth, tenth, fifteenth, eighteenth, and twentieth (lowest). It also gives the scores for the two university groups. Data from Table II are charted in Figure 1.

TABLE II. MEDIANS, QUARTILES, AND RANGES OF SCORES MADE BY VARIOUS SCHOOL GROUPS

	Lowest	First Quartile	Median	Third Quartile	Highest	Cases
U_2	32	60.8	68.6	76.2	94	158
U_1	39	62.8	66.9	71.8	85	57
H. S.—Highest	33	61.2	67.5	74.5	90	38
Second	33	51.0	64.6	71.0	79	24
Fifth	45	54.5	60.5	65.0	74	14
Tenth	25	46.5	56.1	65.1	81	31
Fifteenth	21	32.0	49.5	53.5	69	27
Eighteenth	17	31.0	45.0	62.0	74	32
Lowest	16	26.0	36.0	49.0	57	8
All H. S.	5	40.5	53.8	64.4	91	731

[2] Score as used in this paper means total number of items done correctly.

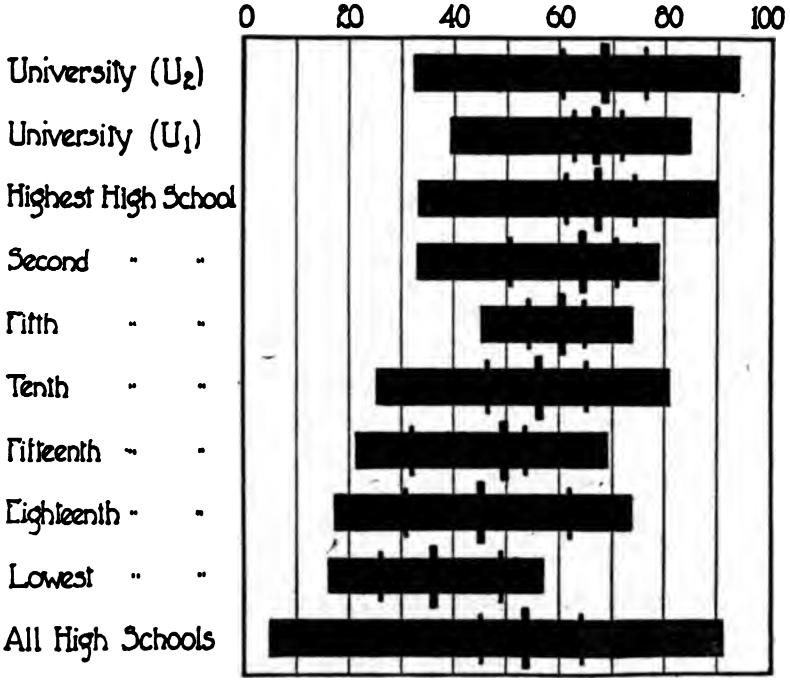


FIGURE I. MEDIAN, QUARTILES, AND RANGES OF SCORES MADE BY VARIOUS SCHOOL GROUPS. (CHARTED FROM DATA IN TABLE II.)

It is clear at once that there is an enormous overlapping of high-school and university-freshmen scores. Note that the median score for the highest high school is actually higher than the median score for the university freshmen who had not had high-school chemistry. Note also that the second high school scored nearly as high as this university group and that down even to the fifth high school the differences in the median scores is inconsiderable.

Table III shows the average percent of scores made on the eight divisions of the test by the same schools as are reported in Table II. Notice in this table that the scores for the highest high school on the fourth and seventh divisions of the test is higher than for either of the university groups. In the first, fourth,

TABLE III. AVERAGE PERCENT SCORES ON EACH OF THE EIGHT DIVISIONS OF THE TEST

	1	2	3	4	5	6	7	8	Cases
U_2..........	86.4	67.2	95.1	34.9	78.6	60.0	44.9	70.8	168
U_1..........	83.2	51.2	95.8	54.8	78.6	52.5	33.7	65.2	157
H.S.–Highest	84.2	43.7	82.0	73.6	80.5	53.0	50.9	63.2	38
Second.....	77.0	57.0	89.0	10.0	76.4	56.2	44.2	66.6	24
Fifth.......	75.7	62.9	95.3	8.6	74.6	57.1	31.4	52.8	14
Tenth......	70.3	37.4	72.0	54.0	76.1	43.5	25.2	36.0	31
Fifteenth...	72.2	40.7	51.6	8.8	68.5	39.6	24.8	48.6	27
Eighteenth..	71.6	34.0	58.0	30.0	69.3	30.6	19.4	31.2	32
Lowest.....	61.3	25.0	46.2	15.0	53.1	33.7	10.0	7.5	8
Average 20 H.S....	70.5	44.0	73.7	35.2	68.8	42.3	27.3	45.6	731

fifth, and seventh divisions, the highest high-school scores are higher than for the U_1 group. The university groups score considerably higher on all divisions of the test than the average of all high schools. The university group which had had high-school chemistry scored higher than the other university group on six of the eight divisions of the test.

Table IV shows the percent of all high-school students who did correctly each of the tasks in divisions 1, 2, 5, 6, and 7 of the test.[3] It also shows the percent correct for the students in the high school which scored highest, and the percent correct for the students in each university group. This table is read: 60.5 percent of all high-school students, 97.3 percent of the students in the high school which scored highest, 89.5 percent of the U_1 group, and 60.0 percent of the U_2 group wrote correctly the name for HgO. This table makes possible a detailed comparison of the differences in the scores made by these four groups. The differences on nearly all items are of the same order of magnitude. There are scarcely any items in the test which

---

[3] Data from the third, fourth, and eighth divisions were not adapted for such analysis as is given in this table. These divisions are therefore omitted.

TABLE IV. PERCENT OF STUDENTS IN EACH GROUP WHO DID EACH TASK CORRECTLY IN DIVISIONS 1, 2, 5, 6, AND 7

Tasks in Divisions 1, 2, 5, 6, and 7	Percent Answering Correctly			
	Total High Schools	Highest High School	U_1	U_2
1. Write the name of compounds from given formula				
HgO	60.5	97.3	89.5	60.0
MnO_2	86.6	97.3	98.2	97.6
$CuSO_4$	93.1	100.0	100.0	98.2
$ZnCl_2$	97.7	100.0	100.0	98.2
NaOH	98.1	89.5[a]	96.5	98.2
$AgNO_3$	89.2	100.0	100.0	98.8
K_2SO_3	39.7	63.1	86.0	82.7
$SbCl_3$	47.8	57.9	59.6	90.0
H_2SO_3	69.8	81.6	58.0	86.9
$KClO_3$	22.0	55.2	44.0	53.5
Average	70.5	84.2	83.2	86.4
2. Write formulas from names of substances				
potassium carbonate	48.6	31.6	66.6	76.2
ferric chloride	38.1	21.0	15.8	80.0
sodium sulfate	60.0	50.0	75.4	85.1
calcium oxide	84.0	79.0	87.6	90.6
silver sulfide	21.9	15.8	28.0	45.3
potassium iodide	78.5	89.5	94.7	97.6
stannic chloride	9.0	15.8	12.2	25.6
ferrous chloride	37.9	39.5	40.3	78.0
chlorine	49.3	79.0	84.2	48.2
arsenious sulfide	12.3	15.8	7.0	45.3
Average	44.0	43.7	51.2	67.2
5. Classification of elements, compounds, and mixtures				
cane sugar	78.2	84.2	96.5	88.1
table salt	82.5	89.5	100.0	95.2
iron	90.1	100.0	96.5	82.1
brass	43.5	65.8	61.4	72.0
quick-lime	76.4	73.1	94.7	90.0
coal gas	52.7	76.3	33.3	51.9
air	79.3	94.7	96.5	86.3
ink	67.8	84.2	72.0	67.8
plaster of paris	70.5	79.0	42.1	68.4
ether	58.9	76.3	93.0	86.3
baking soda	78.9	76.3	98.2	91.8
ammonia	79.5	89.5	89.5	76.2
arsenic	72.5	86.9	93.0	95.2
carbon	88.3	100.0	100.0	96.4
graphite	52.0	81.6	70.1	67.2
glass	42.0	42.1	17.5	34.5
acetylene	64.7	76.3	93.0	87.5
radium	86.4	100.0	94.7	97.0
paint	84.6	92.1	79.0	87.5
steel	26.2	42.1	50.8	50.7
Average	68.8	80.5	78.6	78.6

[a] The percents for the "Highest High School" are occasionally less than those for all high schools combined. For example, in writing the name of the compound NaOH, the general percent for high schools was 98.1, while the percent for the school which in general did better than any other (and which is therefore called the Highest High School) was only 89.5.

TABLE IV (*Continued*)

TASKS IN DIVISIONS 1, 2, 5, 6, and 7	PERCENT ANSWERING CORRECTLY			
	Total High Schools	Highest High School	U_1	U_2
6. COMPLETE EQUATIONS				
HgO (heated)	54.0	87.5	86.0	75.6
$Zn+H_2SO_4$	83.9	97.3	100.0	97.6
$CuO+H_2$	83.5	84.2	94.7	87.5
$Cu+AgNO_3$	24.0	18.4	22.8	44.7
$Fe+Cl_2$	5.4	5.3	0.0	9.5
$KOH+H_2SO_4$	63.5	79.0	80.7	93.4
$Pb(NO_3)_2+H_2S$	70.1	81.6	75.4	91.8
$Cu+H_2SO_4$	12.8	36.8	15.8	27.4
$H_2S+H_2S_2$	16.4	10.6	31.6	33.9
$NH_3+H_2SO_4$	9.4	29.6	17.5	38.7
Average	42.3	53.0	52.4	60.0
7. WRITE THE EQUATIONS FOR GIVEN CHEMICAL CHANGES				
The preparation of oxygen from potassium chlorate	33.3	63.2	31.6	32.7
Burning of hydrogen in air	40.6	79.0	65.0	49.4
Action of sodium upon water	37.1	76.3	52.5	55.3
Preparation of chlorine from manganese dioxide and hydrochloric acid	15.5	49.7	3.5	16.1
Preparation of hydrogen chloride from sulfuric acid and salt	58.5	73.1	77.2	85.1
Test for carbon dioxide	21.2	28.9	22.8	29.4
Laboratory preparation of hydrogen sulfide	24.3	65.8	24.6	53.1
Test for sulfate radical	13.5	28.9	24.6	60.6
Laboratory preparation of ammonia using ammonium chloride	8.7	14.8	12.3	19.6
The action of sulfuric acid on sodium sulfite	20.2	28.9	22.8	47.6
Average	27.3	50.9	33.7	44.9

are especially easy for one group and especially difficult for another. The correlation coefficient (Pearson formula) derived by correlating the percent scores for all high-school students with the percent scores for the students in the U_1 group (column 1 with column 3, Table IV) is 0.84 ± 0.03. The order of difficulty of the tasks for the high-school and the university students is therefore nearly the same.

Tables II, III, and IV show that the difference in the ability of high-school and university students to do the tasks included in this test is not very great. They show in fact that the students in at least one high school excel the students who have had their training in university classes.[4]

COMPARISON OF DIFFERENT HIGH SCHOOLS

This material makes it possible to compare the ability of students who have received their training in small high schools with the ability of those who have received their training in larger schools. Students in large schools seem to do somewhat better. Table V shows the relation between size of school measured in terms of number of students taking chemistry and the median scores. The table is read: one school enrolling between 6 and 10 students in chemistry made a median score between 36 and 40; one made a median score between 51 and 55; and one made a score between 56 and 60. The correlation (rank-difference method) between the number of students enrolled in chemistry and the median scores for the different schools is + 0.58.

TABLE V. RELATION BETWEEN NUMBER OF STUDENTS ENROLLED IN CHEMISTRY AND MEDIAN SCORES FOR ALL HIGH SCHOOLS TAKING THE TEST

Number of Students	Median Scores							
	36-40	41-45	46-50	51-55	56-60	61-65	66-67.5	Total
6-10....	1			1	1			3
11-15....		1	1		1			3
16-20....				2				2
21-25....				1		1		2
26-30....			1		1			2
31-40....		1			1	1	1	4
41-171...		1			3			4
Total..	1	3	2	4	7	2	1	20

---

[4] A study, not yet completed, of ability of students in a selected list of high schools in large cities to do these tasks seems to indicate that students in these selected schools quite commonly excel the students who have received their training in university freshmen classes.

The statement is sometimes made that students enrolled in smaller classes derive an advantage. No support is found in this study for such a statement. Only two of the schools with 15 students or less in chemistry scored up to the mode. Eight of the fifteen schools with more than 15 students in chemistry scored at or above the mode. Eleven of the twenty schools scored between 51 and 60.

The order of difficulty of the tasks in the first, second, fifth, sixth, and seventh divisions for the students in large and small high schools is nearly the same. The percent of the students in the four largest high schools and the percent of the students in the remaining sixteen schools which did each task correctly were separately calculated. The correlation coefficient (Pearson formula) for the relation between the scores made by the students in the four large high schools and the scores made by the students in the sixteen remaining schools is 0.89 ± 0.02.

ADAPTABILITY OF SUBJECT MATTER

Students commonly believe that chemistry is a difficult subject. Studies seem to show that this tradition has become so well established that students of poor ability have a decided tendency to avoid it. This study makes it possible to locate some of the points which cause difficulty. Table III shows that the scores on the second, sixth, and seventh divisions of the test were consistently low for all groups. The second division is a test of ability to write the formulas for given substances; the sixth is a test of ability to complete equations when the right-hand member is given; and the seventh is a test of ability to write complete equations for given chemical changes. Barely half of the responses made to these tasks by the students in the high school which scored the highest were correct. The average score for all high-school students on these three divisions is well under 50 percent. Of these three divisions the seventh is considerably the most difficult. The average score on this division for all students is scarcely more than 25 percent and the average score for the students in the lowest school is but 10 percent.

Table IV shows in detail the items which caused difficulty. In the second division the formulas for calcium oxide and potassium iodide were the easiest ones to write. In each of these compounds the valence of the positive and negative radicals are equal. The valence of the positive and negative radicals of each of the seven other compounds are unequal. The high-school student should be much more familiar with sodium sulfate than with potassium iodide yet a much smaller percent of students are able to write its formula correctly. From such analysis as the foregoing it is possible to make a definite diagnosis of some of the difficulties which high-school students have with chemistry. The low scores for stannic chloride and arsenious sulfide are no doubt due to the fact that these substances are rather unusual in the high-school laboratory.

The difficulties in the sixth division may be located in a similar manner. In the second item ($Zn + H_2 SO_4$) the student is only required to put Zn in place of H_2 and the equation is correct. Similarly in the third item he is only required to put H_2 in place of Cu. All other tasks in this division involve some complications. The low score on the eighth item ($Cu + H_2 SO_4$) cannot be explained by saying that it is unusual. Nearly all students have had the reaction of this equation as a part of laboratory work and it is described in all textbooks. Since seven-eighths of the students were unable to do this task it is clear either that the instruction on this item is poorly given or that the item is ill-adapted to the ability of high-school students. In a similar way in the seventh division it would seem either that most of the instruction covering such items as these is poorly done or that the items are ill-adapted to the ability of the students. The scores for all groups on certain items of the sixth and seventh divisions are so low that they furnish evidence in support of the second of these hypotheses. If this second hypothesis is a fact these data furnish an excellent scientific basis for the revision of high-school textbooks.[5] The teacher with these data before him must choose one of two possible courses: either he must increase the emphasis on these tasks in his instruction or he must eliminate them from his course. Certainly there can be no justification for giving instruction in any material when upwards of 80 or 90 percent of the students are failing to learn it.

---

[5] The procedure for establishing the one or the other of these hypotheses is clear and such investigation is now under way.

The high-school students made very satisfactory scores on the first, third, and fifth divisions. These divisions were respectively, (1) ability to write the names for given formulas, (3) ability to construct formulas from given radicals and their valence, and (5) ability to classify elements, mixtures, and compounds. Those items in the first division which were missed by the most students are somewhat unusual and in some cases of the nature of catch questions. In the fifth division the most common error was that of classifying mixtures as compounds. Probably the largest value of this test material is its value for testing the adaptability of the subject-matter. These data are computed from the results of testing 731 high-school students. It is safe to conclude therefore that these findings are fairly representative of what may be expected from the types of schools in which these students were enrolled.

CONCLUSIONS

1. The difference between students who have taken chemistry in university freshman classes and those who have taken chemistry in the better high schools in ability to do the tasks included in this test is small.

2. Students who have taken chemistry in large high schools do considerably better on the items included in this test than those of the small high schools.

3. High-school students score comparatively high in ability to write the names of compounds from given formulas (division 1); to write formulas from given radicals and their valence (division 3); and to classify elements, mixtures, and compounds (division 5). The lowest scores were on the seventh division which is a test of ability to write the complete equation for given chemical changes. Scores on the second division, which is a test

of ability to write the formula for given substances, and on the sixth division, which is a test of ability to complete equations, were also low.

4. The high-school scores for each of the items of the first, second, fifth, sixth, and seventh divisions of the test are based upon such a large number of cases that the scores on these items may be considered fairly reliable standards for schools of the types represented.

Editorials

If educational service bureaus are to retain, and in some cases regain, the good will of teachers and the respect of superintendents, it is imperative that they subordinate testing to other more vital problems of education. There has been need for the formulation and administration of standard educational and mental tests which without doubt will continue to enjoy increasing usefulness. But well-informed and sympathetic school people are increasingly apprehensive of the danger to the intangible spiritual values of education from over-emphasis upon the testing of the mechanical content of subjects.

Educational tests are rapidly becoming a most potent formalizing factor in public education. They lead teachers to think that success is judged by the use of standard tests on the mechanical phases of instruction, and when teachers think this they necessarily emphasize those aspects of instruction which yield promotion and recognition.

The cure is to be found, not in the abolition of tests by educational service bureaus, but in the development of other interests which will overshadow this essentially minor agency of instruction.

These new major interests should lie close to the firing line. Teachers need help in such activities as methods of instruction, the collection of materials, and the formulation of curricula. Teachers will be greatly aided by difficulty analyses of discipline, and of the teaching of subjects, projects, and ideals. They are clamoring for expert methods of handling these practical difficulties and those who collect them to distribute to the teaching force will earn appreciation and respect.

The time has arrived for the Educational Research Association to put on a broad and vigorous campaign to subordinate the discovery of how well a teacher teaches to the task of helping him to teach better. It is the function of tests to reveal defects; but it is the duty of service to correct them.

W. W. C.

A CHALLENGE

While many research workers are dealing with the minutiae of measurement, the school people are confronting larger problems and seeking light upon them. It is possible so to magnify the importance of detail—the fine points of method and the calibration of instruments—as to lose sight of vital and fundamental issues. Individual differences exist among research workers as in every human group. Some are purists, enslaved by the rigid requirements of a hypertrophied technique. If by chance they stray beyond the orbit by which they are habitually circumscribed, they are afraid they are about to fall off the edge of the world.

To such as these State Superintendent Wood of California administers, though quite unintentionally, a stinging rebuke. We are privileged to present in another part of this issue of the Journal his admirable address given last May before the Conference on Educational Research and Guidance at the San Jose State Teachers College. We are glad to bring his message to the attention of a larger audience than that which heard him on that occasion. For what he says is indeed timely. It is a challenge and an inspiration.

For example, California has a teacher-tenure law; and according to this law the only way the schools may be relieved of weak teachers is to prove their incompetency. Such laws are being passed in many places. The larger cities of the country where many teachers combine to secure legislation have such laws. The theory behind these laws is sound. They are faulty, however, in their administration. The trouble lies in the fact that incompetency cannot be proved. Superintendent Wood says the research workers must devise a method of rating teachers that will "stand up." The time will come when the teacher-tenure laws will be repealed unless they can be administered so as to protect the schools from the incubus of inefficient teachers.

In one city employing about fifteen thousand teachers it is reported that in a certain year only thirteen teachers were dropped from the service because they were rated unsatisfactory. It is incredible that among fifteen thousand teachers less than one-tenth of one percent should be incompetent. What effect this practically universal approval of the services of teachers has upon the morale of the force can only be estimated. That a large number of teachers put forth unusual effort when their colleagues may drift through the year with the same assurance of position and salary is indeed a tribute to the teaching profession.

We have not meant to draw the reader's attention particularly to the rating of teachers as an issue in educational research. There are hundreds of other major problems. Some of them, such as class size, length of term and length of school day, consolidated versus one-room rural schools, are mentioned by Superintendent Wood. He could easily have made his list much longer.

While, therefore, it is entirely proper that we should refine our methods, we must recognize the call for larger vision. We must not hesitate to grapple with problems which demand solution. Appreciating the limitations of our instruments, we may nevertheless rest assured that they are far better than those which are in general use. Consider, for example, the futility of the recent investigation on the part of the teachers of a certain city concerning the value of the all-year-round school. The only real conclusion which this so-called investigation warrants is that the teachers do not want to teach in the summer. Much is said about "social and economic considerations" and the fact that they preclude effective school work in the summer months; but everyone who is so unfortunate as to have to contribute to the work of the world in the summer time knows how fraudulent these representations are.

Counsels of perfection, therefore, on the part of research workers, extended to the "control of all the variables" and to the third decimal place of reliability determinations, may be pressed too far. The decision in regard to grave issues would then be left to those whose lack of any scientfic method, to say nothing of their personal bias, will make them totally unfit for the task. It is well enough in the proper place to be meticulously careful; but to strain at gnats while swallowing camels is proverbially foolish.

B. R. B.

Reviews and Abstracts

BOBBITT, FRANKLIN. *Curriculum-making in Los Angeles.* (Supplementary Educational Monographs, No. 20, June, 1922.) Chicago: University of Chicago Press, 1922. 106 pp.

Professor Bobbitt and the high-school teachers of Los Angeles have undertaken a gigantic task in re-examining the curriculum according to methods described in the monograph "Curriculum-Making in Los Angeles."

To the work Professor Bobbitt brought a comprehensive list of over five hundred "human abilities and characteristics" which was prepared with great patience in his classes in University of Chicago by several hundred experienced teachers who worked under his direction. This list was presented to committees of the twelve hundred junior and senior high-school teachers of Los Angeles who used them as a starting point to determine the "human abilities and characteristics which appear to be generally needed by citizens of Los Angeles." Committees formulated a list of objectives for each of the following subjects and groups of subjects: the social studies (history, social geography, economics, and civics), mathematics, the biological sciences, the physical sciences, physical training, home occupations, practical acts for boys, visual art, music, English expression, modern languages, and Latin. To illustrate, the objectives of musical training finally arrived at are: (1) a mind enriched with the auditory imagery of the great music of the world; (2) ability—as auditor or musical "consumer"—judiciously to utilize music for healthful, abundant, and varied awakening of one's emotional nature; (3) ability, disposition, and habit of participating in home, chorus, or community singing. Note: For those relatively few who go on to higher levels of production, other objectives should be added to the above.

When these have been determined, a set of general assumptions or hypotheses is utilized in formulating the pupil activities by which the objectives are to be realized. Specifically, in music there are twenty-eight hypotheses which will guide the music teachers of Los Angeles in the teaching of music. Some of the hypotheses are: "(1) For the sake of normal training for appreciation, music of proper quality should enter abundantly into the general life of the school. (2) Home player-pianos and phonographs will be utilized for training in musical appreciation as fully as practicable. (3) In the main, for non-performers, music is a thing to be *experienced*, not a thing to be *studied*. (4) Pupils should be made reasonably familiar—so far as ability permits—with the criteria to be employed in judging the worth of music. (5) The Seashore tests should be used in classifying pupils for musical training. (6) The fundamental experience in learning to sing is, of course, to sing. A maximum of time should be given to this with accessory training reduced to the necessary minimum."

The monograph thus deals with two problems—the selection of objectives for subjects and the principles of selection of means by which the objectives are to be accomplished. It does not discuss the resulting content of the subjects.

This study represents a serious and significant attack upon the problems of curriculum construction and is a distinct contribution to the field in two respects. It presents by long odds the most complete sets of objectives and methods of realizing them that have yet been published; and it illustrates and develops consensus as a tool in curriculum construction.

A careful study of the investigation prompts this observation. The venture bulks so large and appears so complicated that it constitutes a strong argument for great research agencies which have ample time, means, and facilities to devote to such studies of the curriculum. It is impracticable for voluntary organizations of teachers in service to carry the task much beyond the point to which it has already been taken.

W. W. CHARTERS

Carnegie Institute of Technology

THOMAS, F. W. *Training for effective study*. New York: Houghton Mifflin Company, 1922. 246 pp.

In this volume, as is pointed out by the editor of the series in which it appears, the emphasis is shifted "from the recitation itself to the preparation for the recitation, from the accumulation of knowledge to learning how to find and use knowledge, from the gathering of information to learning how to use it and hence stand on one's own feet, from drill to appreciation and expression, from learning facts to fitting for responsibilities, and from discipline by rules to training for rational self-control" (p. VIII). It is a mode of approach which, in the judgment of the editor, "could displace, with advantage, much of the special-methods work now given to intending teachers" (p. IX).

The author approaches his task in the conviction that our present-day conceptions of teaching have not, as yet, effected an appropriate revision of methods of studying. We emphasize the need of cultivating independent thinking, but we neglect the application of this view to the process of preparing for the recitation. After pointing out how wrong teaching, the traditional type of examination, and slovenliness in making assignments may foster slavish dependence upon the textbook, the author proceeds to the discussion of the remedies. These fall under the following chapter heads: knowing what is to be studied, having a motive for study, securing continued application, and effective methods of study, such as memorizing, drill work, collecting and organizing information, and solving problems. The final chapter is devoted to the supervision of study, and discusses the physical conditions of study, the use of reference books, the correlation between recitation and study supervision, and other pertinent topics. Each chapter is concluded with a set of suggestive questions for study.

The book is a real contribution. It is clear and simple in style, and offers much useful material. The author knows his ground and handles his subject effectively. Perhaps the most serious defect of the volume is its treatment of problem-solving. As the author himself says, "In a broad sense, the greater part of school work should be connected with the solving of problems" (p. 189). Yet this topic, strangely enough, is dismissed with six scant and formal pages, and the topic of collecting and organizing information is considered without any intimate reference to the solving of problems. It is difficult to see how the aim of "growth in independent self-direction and skill in study" is to bear proper fruit unless more attention is given to this topic, which lies somewhere near the center of the whole subject.

B. H. BODE

Ohio State University

KLINE, LINUS WARD, and CAREY, GERTRUDE L. *A measuring scale for freehand drawing. Part I: Representation.* Baltimore: The Johns Hopkins Press, 1922. 61 pp.

Part I of "A Measuring Scale for Freehand Drawing" by Linus Ward Kline and Gertrude L. Carey, edited by Edward Buchner (being number five of a series of Studies in Education from the press of the Johns Hopkins University), covers the subject of representative drawing. It is an interesting review of a careful study of children's drawings gathered from schools in various parts of the country. These drawings were made under certain conditions on four designated themes for the purpose of establishing a fixed standard by which the progress in pupils' work might be systematically graded.

After a careful analysis, comparison, and grading of the many drawings by unbiased committees, typical drawings were selected and arranged as suggesting a logical development of pupils' abilities in such work. It contains a number of tables of interest to students of a psychological bent.

The report is also illustrated with reproductions of from sixteen to twenty of the selected drawings on each theme. These are so arranged as to offer a key for the scaling of pupil's work.

The resulting scale is intended for use by the pupil as well as the teacher. When used by pupils, it should make for positive growth and development especially along the lines of analysis, perception, and execution. The teacher will find the scale of value as a standard by which to gauge the success of her pupils either during successive periods of the same year or from grade to grade. The promised publication of the reports on design and color will be looked forward to with interest.

WILLIAM H. VOGEL

Director, Department of Art, Cincinnati Public Schools

SNEDDEN, DAVID. *Civic education.* Yonkers-on-Hudson: World Book Company, 1922. 333 pp.

The problem of "making democracy safe," which grew out of the turmoil of the World War and out of the world-wide propaganda for revolution, has centered attention strongly upon the need for better training for citizenship. Professor Snedden's book, "Civic Education," is by far the most thoroughgoing and scientific of all the discussions that have grown out of this crisis. In it he brings to bear upon this important question the analytic methods he has made familiar to us in other applications, but gives many more constructive suggestions than he customarily has done.

In Part I he gives concrete and practical "suggestions to teachers" on the objectives of civic education and materials and methods for attaining them. In Part II he retraces some of the same ground but discusses in a more technical way the sociological meaning of education, the meaning of social education, society's need of civic education, the objectives of civic education, and education for democracy. In Part III he discusses chiefly the problem of methods for finding the objectives of civic education. He favors the method of "type-group" analysis, with which the readers of his "Sociological Determination of Educational Objectives" and his recent journal articles are already familiar. As examples Professor Snedden gives eleven representative type-group studies made by his seminar. These not only illustrate the method but are themselves suggestive through their content. He also has a chapter which gives teachers some very sane guiding principles on the knotty problem of "freedom of teaching in the social sciences" (reprinted from *School and Society*).

The book does not expose some nostrum with the claim that it is a panacea for all our ills. Primarily, it is a challenge to clear and definite thinking about the objectives of civic education and about the means and methods for realizing these objectives. The author would have the educator determine, by sociological methods, what are the threatened civic deficiencies of particular groups of individuals, translate these into very specific educational objectives, and then, unfettered by tradition, search out procedures that would forestall these threatened deficiencies. He expects comparatively little from accidental transfers from school discipline, general moral precepts, or even from history, but much from developmental readings, problems, and projects that are so controlled as to clarify ideas and motivate acts and attitudes specifically related to citizenship.

The major theses of the book are not new to close readers of Professor Snedden's works. However, they are worked out with a wealth of detailed suggestions that justifies anyone, even though already familiar with the view-points, in following them through. To those teachers of the social sciences who are not familiar with Professor Snedden's work the book should come as a revelation. Its sanity, definiteness, and hard-headed common sense are in marked contrast with the vagueness and utopianism of most discussions of the subject. The reader must not, however, expect concrete accounts of method devices. These are left to the teacher to work out. For example, in a rather large section devoted to the possibilities of teaching through problems and projects, long lists of possible exercises are proposed but the actual working out of no one of them is described.

The following seem to the reviewer to be weak points in the volume.

1. It is somewhat fragmentary in make-up. It gives one the impression of having been thrown together from materials prepared for other purposes. Indeed, a number of sections are taken almost bodily from Professor Snedden's other published works. This fragmentariness makes it more difficult to follow through than if it were conceived and executed as a unit.

2. One is embarrassed by the very wealth of its suggestions and problems. It rushes on from one problem to another with such rapidity as to overwhelm one. The book might have been made more valuable to ninety percent of the teachers who will read it if fewer things had been said and these few developed at greater length.

3. While the method of prognoses of type-group deficiencies (made by from one to five judges uncontrolled by any count of cases) is a big step in advance of the crude thinking to which we have been accustomed, it can not be considered a really scientific method. It opens the gate too widely to the easy generalizer. Methods of research in this field must—and can—be worked out that are more objective, more systematic, and more successful in eliminating the personal equation of the investigator.

CHARLES C. PETERS

Ohio Wesleyan

RHODES, CHARLES E. *Effective expression.* Chicago: Gregg Publishing Company, 1921. 532 pp.

A good many teachers are of the opinion that the study and practice of English would not suffer in the least if the publication of books on rhetoric and composition should be prohibited for a decade. Such a damming of the stream is not likely to occur; and it is only fair to say that Charles E. Rhodes goes far toward justifying his own contribution in the form of a volume called "Effective Expression."

Here is a text planned to cover the four years of the high-school course and perhaps the first year of college work. Its range is unusually complete, its arrangement sensible, its material almost superabundant. If one should say that the book contains little which might not be found elsewhere, one would say only what applies to almost every text on high-school and college composition published since Barrett Wendall wrote his English Composition, or since Scott and Denney broke ground for the study of the paragraph.

What "Effective Expression" accomplishes is the presentation, in a way both interesting and manageable, of almost all the principles and forms of expression, written and oral. The collection and organization of material is carefully attended to, with the teacher in mind as well as the student. Mr. Rhodes is old-fashioned enough to believe that the basic principles of rhetoric must be mastered by the student of composition, and has a good deal to say of those abiding qualities, clearness, force, and elegance. His chapter on diction will commend itself to the teacher who must labor with the awkwardness and vulgarity that beset the boy and girl in written and oral work alike. This section is one of the best, admirably direct and simple, yet remarkably complete in its survey.

Another excellent chapter is that on letter writing. It manages to include a brief historical introduction, a comprehensive study of letter forms, plenty of models, and an extended list of exercises.

The study of the paragraph is conventional in treatment, but presents excellent illustrative material. The chapter on oral composition would gain by elaboration of such a helpful section as that on kinds of speeches, and a briefer treatment of matters of poise, vocalization, modulation, and the like, regarding which an observant teacher could easily guide and correct pupils.

There is a lively discussion of narrative writing with especially fresh material relating to dramatization of poems and stories and to the preparation of pageants. High-school teachers will welcome the helps supplied in this section, as well as the suggestive exercises at the end of the chapters on narration and description.

"Effective Expression" presents an adequate treatment of exposition and argumentation, handled in accepted fashion. The chapter on exposition emphasizes notably such special forms of this kind of writing as the exposition of a story and the exposition of character; it contains an interesting section on book reviewing, with plenty of examples; and it suggests helpful plans for the expository composition. Under argumentation one finds serviceable discussion of the debate, with extended explanation of the brief, and many exercises.

Finally, there appears a useful compendium of material on the composition of verse forms.

W. L. GRAVES

Ohio State University

MORRISON, JOHN CAYCE. *The legal status of the city school superintendent.* Baltimore: Warwick and York, 1922. 162 pp.

A good piece of research work always commands the attention of serious students, and to some extent, also, that of the general public. On the basis of this truism one may without hazard predict an extended sale of "The Legal Status of the City School Superintendent" by John Cayce Morrison. It adduces a vast amount of evidence bearing upon a phase of educational organization known to be in a confused state, and shows the extent to which the confusion really exists. The book also brings to light certain things pertaining to the office of city superintendent that seem to be working in the most satisfactory way.

The evidence gathered and subjected to analysis in the course of Mr. Morrison's study comes from four general sources: (1) historical writings; (2) statutes of the various states, and city charters and codes; (3) judgments of lay and professional leaders in education; and (4) general principles of government. The first of these yields relatively little on so modern a phase of educational development. The second constitutes the author's main reliance. The third is considerably developed, although to no great profit, considering the lack of any marked degree of consensus among the replies regarding most of the matters submitted, and the inconclusive value of such consensus in so far as it does obtain. The chapter, dealing with the last type of evidence, is in reality a well-reasoned argument in support of the author's conclusions.

Mr. Morrison's study represents a vast searching of the laws, and his results in this field are of outstanding value. Some of the more noteworthy conditions here revealed are: the practical disappearance of ward representation in boards of education; the predominance of small boards of from five to seven members, and the election by popular vote for a four-year term of office; an almost universal vesting of the boards of education, either expressly or by implication, with the power both to elect and to remove the superintendent; and a prevailing term of office for the superintendent of four years. The study shows that the general advisory function of the superintendent is a matter of established custom rather than of express legal provision. The powers with respect to which the superintendent is most nearly independent of board control, whether in cities operating under general legislation, or otherwise, are: the enforcement of attendance; granting of working papers; the making of formal reports; and the supervision of classroom instruction. The extent, however, to which this last power is expressly conferred by law is surprisingly limited. The office of superintendent obtains an exceptionally large measure of recognition in the state code of Delaware. Kentucky is also a rather conspicuous state in this respect. Well-constructed tables, rather basely treated by the printers, however, occur throughout the text. These show the results of combining in various ways the original data, detailed tabulation of which is furnished on supplementary sheets contained in a cover envelope. The code for interpretation of these tables should have been printed on the same sheets as the tables themselves rather than in the book proper. The extent of defects in this study of the office of the superintendent of schools is by no means such as to affect at all seriously the outstanding value of the work.

F. C. LANDSITTELL

Ohio State University

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors. Address all correspondence concerning this department to Doctor E. J. Ashbaugh, Ohio State University, Columbus, Ohio.

John G. Fowlkes, Ph.D. Columbia, has been appointed assistant professor of education at the University of Wisconsin to give lectures on school administration.

Dr. George W. Frasier has resigned his position as director of classification and statistics in the Denver public schools to become dean of the graduate school and professor of education in the Colorado State Teachers' College at Greeley.

Mr. Ray Lathan, who for some years was superintendent at Le Mars, Iowa, has lately been appointed Director of Research at Duluth, Minnesota.

At Lorain, Ohio, all the junior high-school pupils are being classified as far as administrative restrictions permit on the basis of the results of the Illinois Intelligence Scale given last September. Pupils entering the first grade are also being grouped. The Detroit First Grade Intelligence Test is being used for this purpose. Slow, average, and superior groups are being set up both for the junior high-school and first-grade children.

W. H. Hughes moves pretty fast. Appointed a few months ago to head up the Department of Educational Research at Pasadena, California, he is out with Volume I, Number 1 (dated September 26, 1922), of the Educational Research Bulletin. It is dedicated to the encouragement of the spirit and practice of cooperative research. We feel sure that the Pasadena schools will derive much benefit from the Bulletin. Although essentially local, it will no doubt prove of interest to educational students everywhere. We wish the new Bulletin every success.

Dr. Margaret S. McNaught, for nine years Commissioner of Elementary Education for the State of California, has tendered her resignation to the State Board of Education. She plans to spend several years in Europe.

Previous to accepting the commissionership of elementary education, Dr. McNaught, then Dr. Shallenberger, was for several years professor of psychology at the San Jose State Normal School.

Miss Dora B. Thompson, Assistant Superintendent of Schools, Boise, Idaho, submits a plan for use in determining promotions. It is designed to clarify the ideas of teachers concerning the meaning of the marks they give pupils. A chart of fourteen items constituting the factors which, in the opinion of the school officers, should be in the mind of the teacher in rating the pupil is provided. Ratings in mental ability, in achievement, and in knowledge of subject matter are included.

Doctor May Ayres Burgess, formerly in the Education Department of The Russell Sage Foundation, is now with the Motion Picture Producers and Distributors of America of which Will H. Hays is president. Doctor Burgess is trying to act as an informal liaison official between the motion picture industry and education. She writes that quite a bit of research is likely to be involved and hopes that she may have a part in it.

We have just received Survey Bulletin, No. 2, from Superintendent A. W. Elliott of Mount Vernon, Ohio. This bulletin gives for each class in the elementary schools the average intelligence quotients and the scores at the beginning and end of the year in spelling, reading (Monroe), composition (Willing), and arithmetic (Monroe Reasoning). Very gratifying progress was made between the beginning and end of the year in the abilities measured by these tests. Nevertheless, as usual it is found that "there is a great deal of difference in teaching skill."

The Bureau of Naturalization of the United States Department of Labor has issued a pamphlet entitled "Suggestions for Securing and Holding Attendance of Foreign-Born Adults upon Public-School English and Citizenship Classes" which is extremely interesting and suggestive to all interested in the problems of the foreign born. Not only does it give some extremely practical suggestions concerning the subjects indicated by the title, but it also gives some decidedly good hints regarding the methodology of handling the groups. The teaching devices described are those which have proved successful in many good schools, and are further supplemented by reports, from a number of centers in different parts of the country, which suggest practices demonstrated valuable in specific local situations. The pamphlet can be secured by writing to the Bureau of Naturalization, Washington, D. C.

Miss Myrtle L. Kaufmann, elementary supervisor at Logansport, Indiana, has furnished her teachers with an analysis based on the frequency of use of the phonic elements occurring in the primer used at Logansport. In her bulletin on the subject she says: "It seems reasonable to assume that those sounds which occur most frequently in the early reading vocabulary are the ones which should be taught first. In the first place, they are more easily taught because they are the most amply illustrated in the child's practical experience with words. In the second place, the fact that they occur often is a guarantee of their frequent use after they are taught.''

This practical device of Miss Kaufmann's is in line with her study of the frequency of the appearance of words and phrases—a study which appeared in the June, 1922, number of the *Journal of Educational Method.*

In the *School Index* for September 29 we find the following summary of the enrollments of the Cincinnati schools for the present year:

	1922	*1921*	*Increase*
Elementary	38,012	37,740	272
Special	1,534	1,260	274
High	8,136	7,671	465
Total	47,682	46,671	1,011

The New York Society for the Experimental Study of Education is moving to organize the Educational Foundation. The purpose of this foundation is to finance scientific studies in education. A large number of men and women prominent not only in educational but in professional and business life have signified their interest.

It is planned to perfect the organization of the foundation during the present school year. A statement emanating from the experimental society points out the need for the foundation in the following words: ''Decisions on school questions should not be left to chance. Tradition and opinion should give place to carefully observed and verified facts. But adequate facts are not fully available. There is much need for painstaking, scientific research in education, the type of research that has produced such striking results in the natural sciences and industry.''

Miss Nina C. Vandewalker of the Federal Bureau reports that the growth in kindergarten registration from 1918 to 1920 amounted to 37,811 thus raising the percent of kindergarten attendance from ten and one-half percent to eleven. The states which report the largest gains are California, New Jersey, Michigan, Minnesota, Iowa, Massachusetts, and Illinois, in the order named.

Margaret C. Holmes, assistant director of kindergarten in New York City, reports that she and her colleagues are working on two problems as follows: (1) an attempt to measure the relationship between the Detroit Kindergarten Test and the Binet-Simon Test; and (2) an attempt to measure the relationship between the teachers' judgments and the ratings obtained by the use of the Detroit Kindergarten Test. Some work has been done on the first of these problems and the second is well on its way to completion. Ten kindergarten teachers each classified twenty children as advanced, normal, or slow, after an experience with them of about three weeks. The Detroit Test was then given to the same children and they were divided into three groups on the test basis. The agreement between the teachers and the tests in the classification of the children ranged from 78 percent placed in the same group by teacher number I to 40 percent by teacher number X.

Per Pupil Cost Per Recitation in High-School Subjects

Assistant Superintendent Dawson of Grand Rapids has issued a mimeographed bulletin showing the per pupil recitation cost on various high-school subjects. The range is from 3.7 cents for gymnasium to 14.3 cents for public speaking. The average cost for all subjects is approximately 6.7 cents.

He points out that there are, in general, three causes of differences in cost: (1) the salary of the instructor; (2) the number of periods per week taught by the instructor; (3) the size of classes.

The following regulations control program-making in the Grand Rapids high schools.

1. It is undesirable to form classes with less than fifteen students.
2. For schools having sixty-minute periods, regular academic teachers should teach five periods, manual-training teachers should teach six periods, and laboratory work should be counted in the same manner as manual training.
3. Session-room teachers, with rooms having less than one hundred pupils, should teach four or five periods per day; those having one hundred to two hundred pupils should teach three to four periods per day; and those with rooms having more than two hundred pupils should teach two to three periods per day.
4. In schools having forty-five minute periods, regular academic teachers should teach six periods, manual-training and laboratory, eight periods; and session-room teachers, one period more than in schools having sixty-minute periods.

Mr. Dawson reaches the conclusion that these regulations, which are in accord with those of the North Central Association of Colleges and Secondary Schools, are fundamentally correct and that in the administration of the high schools these regulations should be adhered to in program-making.

He also states that an outstanding feature of the report is the very low per pupil per recitation cost of public school work as compared with private instruction. The inference is that parents should think how much less the cost per child is in the public schools compared with cost of private instruction rather than how large the school budget is as a whole.

Reply to a Criticism of Tests Requiring Alternative Responses

In an article in the October, 1922 issue of the Journal, regarding the reliability of tests requiring alternative responses, Dean Hahn has voiced criticisms which many others may be thinking. It seems important, however, that we should not neglect in such a discussion, to consider certain important relevant facts. In the first place, the best test of any test is correlation with a criterion. If this correlation is satisfactory, we can forget all minor criticisms concerning chance, etc. If we accept with Gates that the best available criterion of success in a course is a composite of all the grades given, including daily recitations and oral quizzes, special conferences, written papers, and written examinations of various types, then the fact that the true-false test has a higher correlation with this criterion than the conventional examination must be accepted as significant evidence of the validity of this type of test as a measuring device. But even though one denies this criterion, he must choose some other. Dean Hahn tacitly assumes some criterion of achievement. In contending that the conventional examination is better than the true-false he is assuming for the former, better predictive value as measured against this presumably entirely subjective criterion. Nowadays in the light of the evidence, not many are willing to rely on such subjective standards.

In the second place, no one has a right to insist on perfect reliability. The question at issue is "Does this method secure more accurate results than other available methods?" To answer this, the reliability of the true-false method should be compared with that of the conventional method of examination making. Reliability may be defined as the degree to which a test agrees with itself in measuring the same abilities on two or more occasions. While high reliability does not in itself mean that the test measures the abilities we want to measure, nevertheless the significance of the test in that respect is always restricted by the degree of reliability. The unreliability of the ordinary examination is notorious. Reliability coefficients as low as 0.35 have been found. Whatever abilities it is designed to measure, it is clear that such a test as an index of those abilities is vitiated by the intrusion of irrelevant factors which are not constant. Woods has shown that even when the accuracy of the scoring is unusually high, reliability coefficients of only 0.76 and 0.61 were obtained from three-hour examinations in Algebra and Geometry given by the College Entrance Examination Board. These can hardly be considered satisfactory since other forms of tests exist which are more reliable. The newer type of examination used in Columbia University yields a coefficient of 0.90. This includes several types of examination, among which is the true-false. The latter alone has a reliability of 0.80. Toops has found that of three of the newer types of test, namely recall, recognition, and true-false, the latter yields results as reliable as the others for equal amounts of examination time given.

It is in the attempt to eradicate the elements contributing to unreliability that the true-false tests have come into being along with other new examination methods, and not primarily "to avoid trouble in answering questions and scoring answers," although this is a most valuable outcome and not an unworthy argument in favor of such tests. That they have been successful in the attempt to increase reliability and validity has been proved. That they are perfect indices of ability has not been claimed even by their most ardent advocates.

It is true that with the method of scoring used, one assumes that every wrong answer is a guess. It will be guessing unless (a) the pupil has learned the fact erroneously, or (b) the wording of the question is suggestive. Erroneous knowledge may for practical purposes be considered negligible and as to (b) there is as much reason to believe that the pupil has been suggested in the right direction as often as in the wrong. Nor does the supposition that the number of wrong guesses equals the right guesses always work with precise accuracy. But given a sufficiently large number of questions, the effect of this inaccuracy on the total score is negligible. The power of this type of test, to cover in a given time far more ground than the ordinary test can, adds more to its relative reliability than any such inaccuracies can subtract. The ordinary test suffers from the fact that the smaller the number of questions, the greater the effect of chance vitiations.

That alternative-response tests may be so made that they constitute a test of judgment and thought-organization has been demonstrated by Wood. Inability to test the "why" is a criticism of the skill of the test-maker rather than an indictment of the test form.

The question of the pedagogical value of examinations is a pertinent one. It is always necessary to consider the purpose for which any particular test is given. Many forms of tests may have excellent value as teaching devices but at the same time be unreliable indices of students' relative standings. On the other hand, a measuring device need not of necessity be a teaching device. To insist that every test must have value as a teaching instrument is to deny that the measurement of the results of teaching may be legitimately the primary motive for the giving of an examination. As education is administered today, however, the necessity for reliable measuring instruments is apparent.

It is true that any test involving dichotomous responsing is weak as an instrument for diagnosing special individual difficulties, although for general diagnosis it has proved its usefulness. If a test could be devised which would have all the advantages of the true-false and correct for this disadvantage, it would be heartily welcomed.

H. M. BARTHELMESS

Uses of Tests in State Normal School at Ellensburg, Washington

During the first few weeks after entering, students of the State Normal School at Ellensburg, Washington, are given intelligence tests and standardized tests in reading, arithmetic, grammar, geography, and history. These data were worked up by the department of psychology for the use of all departments. Last year the Army Alpha was given, this year the Thurstone test will also be used. Check tests are used in certain cases and specialized tests when possible.

Last year the upper, middle, and lower quartile were found and students were listed in their respective quartiles for both intelligence and educational tests.

In order to make the data easy to read, red circles were placed around those marks falling in the upper quartile, green circles around those marks falling in the lower quartile, or in some educational tests around those falling below the eighth-grade median which was considered standard for normal-school students.

The upper-quartile group and the lower-quartile group present striking appearances and even to the eye the correlation between intelligence rating and standing in subject matter is obvious. Upper-quartile students show a great predominance of red circles and the lower quartile of green circles.

Those students whose scores in subject matter fall in the lower quartile, or below the eighth- or ninth-grade medians, were assigned to remedial classes or to further work in the subjects in which they were weak.

A definite attempt was made to give those students in the upper quartile opportunity and encouragement for advancement and leadership.

Individuals who showed rather high rating in intelligence but low scores in subject matter were interviewed by the psychologist in an endeavor to get at the reasons for the discrepancy, for purposes of remedying the situation.

Efforts were made to raise the academic standing of those who fell in the lower quartile in both intelligence and subject matter, but little improvement was made. Some of these students were eliminated by repeated failures and others were advised to go into other vocations. It seems that those who rank C and below in the Army Alpha test are not teacher material.

A study of standings in each subject each quarter was also made and curves were plotted of intelligence ratings and markings. Those who were high in intelligence were on the whole rated high in subject matter and those who were low in intelligence were low in subject matter. Some very interesting information was obtained from the teachers' marks; namely, that bright students were working under their ability, average and below-average were working at their best.

As was found in Kirkpatrick's Massachusetts study, we found our teaching directed to the lower and average rather than the upper levels of student ability.

All the scores a student obtains in these examinations are filed permanently on a Teacher Rating Sheet which is used in the appointment office. The result of the physical examination is also placed here. In addition, five qualities thought to be important for teachers are scored by the Scott rating scheme and about twenty points are left for word descriptions.

During the coming year a more elaborate program is contemplated. The high school records will be studied. A personal information card similar to that used at Brown University will be used. The Thurstone as well as the Army Alpha test will be given.

MARY A. GRUPE

Reading Improved Through Tests in Iron Mountain, Michigan

In order to determine the condition of reading in the Iron Mountain public schools, the Burgess Silent Reading Test was given by the supervisors in grades III to VI, inclusive, in January 1921. The results reveal some startling facts. Not only was speed much below standard, but there was a great deficiency in comprehension. In numerous cases children who read with fluency and expression and were considered good readers by the teacher failed completely on this test of comprehension. As a result of a conference between teachers and supervisors the pupils in each room were divided into groups as follows: (1) Group A—those who ranked above standard in both speed and comprehension; (2) Group B—those who ranked as average readers but needed general help; and (3) Group C—those who ranked below in either speed or comprehension or in both.

Group A did very little, if any, oral reading but spent their time in individual silent reading either for pleasure or for information. They were frequently called upon to report upon what they had read. Group B read daily, the emphasis being divided between speed and comprehension as their needs were recognized. Group C was given added time and attention and a careful study was made of their special difficulties. These groups were very flexible and promotions made in them as progress warranted. This gave much inspiration to the work.

Individual teachers instituted excellent methods of remedial work that were most effective. Prominent among these were the following:

1. Suggestions for decreasing vocalization given in O'Brien's "Silent Reading."

2. Silent reading of a given amount under pressure of time followed by reproduction.

3. Questions upon the blackboard to be answered through study. An illustration taken from a third-grade study of Stevenson's "Land of Story Books":

    1. When did it say the lamp was lighted?
    2. Where did it say his parents sit?
    3. Did they sing and talk?
    4. What did he carry with him?
    5. Where did the lions come to drink?
    6. What kind of camp did he play he was in?
    7. Who stopped his play?
    8. Has the little boy really been across the sea?
    9. Where did he go after play?
    10. If you think this little boy liked to play draw a line around play.

4. Pupils frequently prepared their own questions either orally or in writing. Other pupils were called upon to answer these and were encouraged to evaluate the type of question.

5. Studying over a new lesson with the pupils for meaning of words, phrases, idioms, etc., before attempting to read it.

6. Dramatization of interesting stories using characteristic phrases.

7. Continuous analysis of meaning of words through context and use of dictionary. When practicable these words were incorporated into everyday usage.

8. In grade III riddles were composed by teacher and pupils for quick interpretation.

9. Flashing phrases and sentences before pupils for speed in comprehension.

10. Supplementary tests of various kinds were sent from the office and others prepared by teachers frequently to check on individual progress.

This work was not confined to the grades from three to six. The first- and second-grade teachers realized that the real foundation for future good reading rested with them. They studied the problem as intensively as the others, followed many of the above methods, and were confident that the results would be in evidence next year.

At the close of the year, Picture Supplement No. 2 of the Burgess test was given and a careful comparison of results tabulated. The improvement was most gratifying and showed that a forward step was being taken in the right direction.

Eighty-eight percent of the whole number of pupils increased their speed and there was an improvement in speed of 36 percent.

There was a marked improvement in both accuracy and comprehension as the following table shows:

	January Total	June Total
Right	4549	6867
Wrong	1520	1911

Not only do these data prove that definite measurable progress was made but there were many other evidences of improvement. The pupils were able to do more intelligent work in geography and history, use better judgment in selecting material and topics, appreciate literature better, and interpret thought problems in arithmetic much more clearly.

Those directing the work are gratified with the progress made during this short period and expect that with two or three years of this intensive work, reading in the Iron Mountain Schools will have reached a very high standard.[1]

M. B. TRAVIS

Superintendent, Iron Mountain, Michigan

Classification Into Ability Groups In Santa Rosa

Eighty-nine pupils entering junior high school in September 1921 were classified into ability groups on the basis of testing with the National Intelligence Test, Form I, Scales A and B. Those having mental ages one or more years above their chronological ages were placed in VII-A^1 group. Those whose mental ages were one or more years below their chronological ages were placed in VII-A^3 group.

The junior-high teachers were told of the separation into ability groups and it was suggested that more be expected of the VII-A^1 group and less of the VII-A^3 group than of the VII-A^2 group. It was also made clear to the teachers that a recommendation from them for the shifting of pupils from one group to the other would always be considered.

---

[1] We do not wish to discourage Superintendent Travis and his teachers, nor minimize the value of the work which they have done on the reading situation. It is but fair, however, to call the readers' attention to the fact that various investigators have found evidence that the Picture Supplement, No. 2, of the Burgess Test is not as difficult as No. 1. An article by Professor Gates of Columbia University which will appear in an early issue of this Journal gives the results of some very careful work along this line. In other words, it is probable that the improvement in the reading situation in Iron Mountain, while significant, is not so great as the figures obtained would seem to indicate.

EDITOR.

During the term one boy and three girls, on the recommendation of their teachers, were shifted from the original grouping. The boy, who had been placed in the VII-A^1 group, was placed in the VII-A^3 and the two girls were placed in the VII-A^3 from the VII-A^2. It is interesting to note that at the end of the school year two of these three each made a score of 122 in the Terman Group Test. The third had been transferred to another school. However, these pupils seemed to be misfits in the groups in which they were originally placed and so the change was made.

In February, 1922, the VII-A ability groups became VII-B ability groups. There were a few changes in the personnel of the groups owing to transfers from other schools but an effort was made to place all new comers in the group most nearly fitted to their abilities.

In March, 1922, the following tests were used in each of the groups:

1. Terman Group Test.
2. Trabue Language, B and C.
3. Monroe's Silent Reading Test II.
4. Woody-McCall Mixed Fundamentals.
5. Starch's Problems, Scale A.

The median scores obtained were as follows:

	VII-B^1	VII-B^2	VII-B^3
Terman Group.........	104.5	86.5	62:5
Trabue Language (average B and C)....	15.1	14.3	12.7
Woody-McCall.........	29.0	27.5	27.3
Starch Problems........	11.6	10.0	8.9
Monroe's Reading-Rate.	133.0	114.0	108.0
Comprehension.........	35.0	24.3	20.8

In the educational tests the greatest difference in abilities is shown in the reading. Monroe's Test II for Grades VI, VII, and VIII failed to measure the VII-B^1 group and so Monroe's Test III for grades IX, X, XI, XII was used. The score made in this test was so far above the norm for grade VII that it was decided to use this same test with an X-B English class (an unselected group) under as nearly identical conditions as possible. This was done and while the X-B class made a score above the standard for grade X the VII-B group excelled the X-B group in rate 47.7 points, and in comprehension 6.6 points.

In order to check these results, Thorndike's Alpha 2 "Understanding of Sentences" was used and the median score for this VII-B^1 group was equal to Thorndike's standard for grade IX. The results of the tests suggest:

1. That the academic work done by the pupils closely parallel the mental status as indicated by the intelligence tests.

2. If the Woody-McCall norms are not too high, then our seventh grade is weak in the fundamentals of arithmetic.

3. Either the norms of Monroe's Reading Test are too low or our seventh grade has a reading ability which is above the average grade XII student in both rate and comprehension.

4. The type of reading ability that is measured by Monroe has probably been stressed more in our schools than the type measured by Thorndike.

5. The pupils of VII-B^2 and VII-B^3 groups are probably working under conditions fairly well suited to their abilities while in the VII-B^1 group much time is being spent unprofitably.

6. To accommodate the different abilities of our different groups a more elastic program is needed.

M. LOUISE IRVING

Santa Rosa, Cal.

The Retention of Pupils in the Blewett Junior High School

One of the claims made for the junior high school by way of justifying the greater expense, more arduous labor, and disconcerting departure from traditional procedure, incident to the inauguration of the movement, is that the new type of intermediate school will operate to retain its pupils to a degree not accomplished under the former plan. The writer has always been interested in the question of the retention of adolescents in school, and, upon assuming his present position, attempted to discover whether the institution had been living up to its billboards in that respect. He soon found that the question could not be easily answered. In addition to the legitimate snares which attend the use of the statistical devices recommended by Strayer, Thorndike, Ayres, and others for the purpose, there were other complications. The district from which the school was fed had been changed many times. High rents were driving people out of the district. There were three different rates of movement through the course of study, so that pupils who entered with certain other less capable children were later perhaps a year ahead of them in grade. In short, there were so many factors affecting the enrollment at various points that it was impossible to estimate the elimination by comparing the enrollment of the ninth grade with that of the seventh, the enrollment of the ninth with that of the seventh two years ago.

In order that eventually there might be an answer to the question we resorted to the expedient of following the individual pupil from the time of his entrance until he severed his connection with the school in one way or another. A special "Class Registration Card" was devised which enabled the class supervisor to keep check on the individual students through a series of semesters.

When a class of pupils entered from the contributing elementary schools they were divided into advisory groups on the basis of school ability, it being the intention to maintain the integrity of each group throughout the stay of its members in school. Incidentally, it was intended to have the group in charge of the same teacher adviser throughout. The adviser entered the names of the members of the group upon the class register card at once. At the end of each half-year thereafter he made an entry opposite each name to indicate present membership, transfer, withdrawal, etc. A pupil received, by transfer or original entry, was listed at once. These entries were simplified by means of the code of abbreviations.

This began with the class which entered in September, 1920. Now, at the end of two years, we know positively some very definite things about the degree of elimination which has taken place in this school in that time.

In Table I is given a "Statement of Loss and Gain" by way of summary. One half of the table deals with a combination of the two classes which entered in September and November, respectively. These two must be studied together, since they were so close together in grading that they became mingled soon after entrance. Since November, 1920, classes have entered only on the half-year. The other half of the table deals with the sum total of all the classes which have entered the seventh grade from September 1920, to January 1922, inclusive. Both these statements are as of June 1922. Later figures are not available as yet.

TABLE I. STATEMENT OF LOSS AND GAIN

	Classes Entering			
	Sept. 1920–Nov. 1920		Sept. 1920–Jan. 1922	
Original number at entrance		255		947
Transferred out and by death	55		123[a]	
Transferred in	40		90	
Net loss by transfer		15		33
Net receipts		240		914
Withdrawals, financial reasons	3		3	
Withdrawals, failure	4		6[b]	
Withdrawals, health	3		7	
Withdrawals, cause unknown	3		10	
Withdrawals, gross	13		26	
Re-entered after withdrawal	4		5	
Net loss by withdrawal		9		21
Number remaining, June 1922		231		893
Net loss	9/240=3.75%		21/914=2.3%	

[a] Loss of one by death

[b] One, "Don't like school"

The writer was astounded when he saw these summaries. Of the 255 pupils who entered in the fall of 1922, less 55 transferred out, plus 40 additions, only nine have "quit school." Sixty-four are ready for tenth-grade work in the senior high school; 66 for the second half of their ninth-grade work; and 101 to begin ninth-grade work. Similarly, of the 947 pupils who have entered the seventh grade since we began to keep these records, plus 90 additions, only 21 have "quit school." It will be noted that while we do not assume responsibility for the present academic status of those who left us to continue their schooling elsewhere, we do assume that responsibility for those who have been received by transfer. It should be further noted that the net loss through the transfer relationship with other schools is the result of the prevalence of high rents in this district, there having been practically no building in this neighborhood in recent years.

There are two other phases of the elimination question which bear upon the subject. One is the loss in transition from the junior high school to the senior. This we cannot know definitely until this time next year. The other is the loss in transition from the elementary school to the junior high school. Table II gives a "Statement of Loss in Transition" which represents studies of the two last entering classes, those of January 1922 and September 1922, which shows that the loss here is negligible.

TABLE II. STATEMENT OF LOSS IN TRANSITION FROM ELEMENTARY SCHOOLS TO BLEWETT CLASSES ENTERING JANUARY 1922 AND SEPTEMBER 1922

	Class Entering			
	January 1922		September 1922	
Total number finishing elementary schools		229		182
Number not in Blewett district	9		8	
Number who left city	0		14	
Total eligible list		220		160
Number entering private schools		3		5
Net Blewett list		217		155
Number withdrawn because of health	0		2	
Number withdrawn to go to work	1		0	
Number who entered Blewett		216		153

The reader may draw his own conclusions. It would seem from the above that it is entirely possible for the junior high school to eliminate elimination.

H. H. Ryan

Principal, Blewett Junior High School, St. Louis, Missouri

Educational Research Association

(Frank N. Freeman, *Secretary and Editor*)

EDUCATIONAL MEASUREMENTS AT WEST ALLIS

Reported by T. L. Torgerson, Director of Department of Educational Measurements, West Allis

The following extracts are taken from a report of the Department of Educational Measurements of West Allis, Wisconsin. This report, made by Mr. T. L. Torgerson, the director, illustrates the possibilities of such a research department in a small city school system.

***Organization and Administration.*—The local school system is organized on the six-six plan. The kindergarten and first six grades comprise the elementary schools, and the upper six grades the secondary school or junior-senior high school. A Department of Educational Measurements was organized September 1920 with a director and an assistant in charge.**

The results of achievement and intelligence tests supplemented by the teacher's judgment form the basis for promotion in the elementary school. Final examinations begin in the fifth grade. Pupils are classified according to ability. A permanent record is kept of every pupil in the elementary schools giving the results of standard tests. A system of record forms has been devised for this purpose. Every pupil in the elementary and junior high school receives at least one intelligence test every year. The subnormal pupils are segregated and placed under the care of special teachers.

***Classification.*—Each grade in the junior high school is classified according to ability. The intelligence quotient forms the basis for this classification. Thus each section approximates a homogeneous group and the entire class can be taught as one individual. Irregularities in promotion and administrative**

The Table I shows the classification in the junior high school during the past semester.

TABLE I. MEDIAN INTELLIGENCE QUOTIENTS. ILLINOIS GENERAL INTELLIGENCE TEST. FEBRUARY 1922

SECTIONS	BOYS	GIRLS	BOTH	NUMBER PUPILS
VII-B^1	88	88	88	18
VII-B^2	116	111	114	16
VII-A^1	73	78	76	38
VII-A^2	82	88	85	38
VII-A^3	96	96	96	28
VII-A^4	109	109	109	
VII-A^5	120	121	121	

12.2 percent to 7 percent. While failures have decreased the standards of work have been raised.

In the elementary school we have aimed to provide for the superior pupils by classification, maximum assignment, and extra promotions. The pupils are placed in grades suited to their mental ability. The slow pupils are given minimum assignment with more drill. Undoubtedly the best way of meeting the problem of individual differences is by organizing the elementary school on a multiple track basis. This system provides a course of study for the slow pupils, another for the average, and still another for the superior pupils. Consider the advantage of having a course of study fitted to the needs of each group. Failures are eliminated. All of the pupils complete the elementary school in eight years and they are all happy, since failures, misfits, and the resulting eliminations have been done away. The schools of Oakland, California under the direction of Virgil E. Dickson, Director of Research, are organized on this basis.

In the junior high-school readjustment of the curriculum to conform to the abilities of the various groups was given considerable attention during the past year. It is hoped that progress will be made in this direction during the present year.

A program of this kind will cost no more to administer, on the contrary, it will be more economical as failures will be reduced to a minimum and the cost of instructing the pupils who otherwise repeat their work will be eliminated. The pupils will be kept in school longer because they are given the kind of work which they can do and for which they are fitted. In brief the ultimate result of such a program will be:

1. Keeping the slow pupils in school at least through the junior high school.
2. Reducing elimination of all pupils to a minimum.
3. Reducing failures to a minimum.
4. Elimination of retardation.
5. Reducing the cost of instruction by eliminating repeaters.
6. A program of educational and vocational guidance that will determine the fitness of the individual for a given vocation and direct his training toward the vocation for which he is fitted.
7. Training all pupils to be 100 percent efficient.

JOURNAL *of* EDUCATIONAL RESEARCH

Volume VI DECEMBER, 1922 Number 5

PROFESSOR TERMAN'S DETERMINISM: A REJOINDER

WILLIAM C. BAGLEY
Teachers College, Columbia University

In the JOURNAL OF EDUCATIONAL RESEARCH for June, 1922 Professor Terman does me the honor to discuss editorially certain parts of the paper on "Educational Determinism," which I read at the Chicago meeting of the Society of College Teachers of Education.

I have to thank Professor Terman for the clearness with which he has summarized my principal charges against the group of specialists in mental measurements whom I characterized as educational determinists. I am glad to learn that he repudiates the clear implications of the paragraph that I cited from one of his books relative to educational differentiation at the age of twelve on the basis of assumed vocational destinies as determined by the I.Q., and that he disclaims for himself and his fellow-workers the editorial statement that I quoted from one of their technical journals to the effect that the school does not really educate so much as it certifies to educability. I may also be permitted at this point sincerely to reaffirm my warm appreciation of Professor Terman's sterling contributions to the science of education—an appreciation which, I am glad to say, has been publicly expressed on more than one occasion.

It is, therefore, only with the keenest regret that I find myself still at variance with Professor Terman in respect to the far-reaching questions which the subject under discussion involves. I may have "read a meaning into the words of the 'determinist' which was never intended" (Dr. Terman's editorial, p. 60).

The "intention" of the determinists whose writings I quoted I had but one means of judging—namely, the words in which they expressed themselves. I labor under the same limitation in attempting through a study of Professor Terman's editorial to correct what may have been an error and a possible consequent injustice in my earlier paper. Notwithstanding his disclaimers, it seems to me that his brief and "breezy" reply to my rather serious charges fully confirms my earlier conviction that the determinists, because of their hereditarian preconceptions, are grossly misinterpreting the facts which the measurements have revealed.

I do not think, for example, that one who is not pretty thoroughly obsessed by such preconceptions would use a false and misleading analogy with reference to so important a matter as that which is involved in the "right of the individual to share in the spiritual life of the race."[1] It was and is my contention that this right should not be invaded or invalidated short of the most conclusive proof that the individual is quite incompetent to avail himself of his share in the human heritage. Professor Terman accuses me of thus "resting my case on the logical impossibility of proving a 'universal negative.'" And then, utterly to rout me, he points triumphantly to the fact that "no one has ever proved the impossibility of life returning forty-eight hours after death, yet we do not hesitate to bury our dead within that time." Yes, truly; but we do not bury them before they die, although we may be very certain that death is imminent.

Nor will a thoroughly unprejudiced writer be so incautious as to use a metaphor in a misleading way. Not that I object to figures of speech. Indeed, Professor Terman's reference to my paper as an "oratorical bombshell" is positively flattering. Figures of speech are legitimate and often invaluable in controversy as in other forms of discourse; but one ought not to sacrifice truth for the sake of a clever phrase. Professor Terman speaks of my "rhapsodic peroration . . . . . . on the ultimate illumination of the world by gleams of light struck from dull minds." The sentence is a rhetorical achievement—but it is quite the opposite of what I said. It is the light that is *let in* to common minds that will illuminate the world if anything does.

---

[1] I have this phrase from Professor B. H. Bode, who says in a personal letter anent this controversy: "The right of the individual to share in the spiritual life of the race is really the issue. This is *the* important item in any properly constructed program of education."

These are, perhaps, minor matters of disagreement. Not so to be characterized is the fundamental question at issue—namely, the possibility of developing through educative and other environmental agencies the traits that the determinists assume to be both native and essentially unmodifiable.

In my paper I accepted as an hypothesis the probable existence of certain biological traits underlying and conditioning mental growth. I assumed that the structures and functions constituting these biological traits were subject to the biological laws of variation and growth. In respect to their growth, I assumed that they matured fairly early in life. In so far as I can learn, the sum of these traits makes up the hypothetical entity that the determinists call "general intelligence." Whatever mental development may be due exclusively to the natural maturation of the structures and functions presumably underlying general intelligence I termed, for convenience, "vertical growth." With it I contrasted the mental development that comes from the acquisition and organization of experiences, whether gained directly from adjusting to actual situations or indirectly through a mastery of the social heritage. This variety of mental growth I characterized as "horizontal."

I make no claim of originality for drawing this contrast. It is obvious enough once the premises are assumed; hence I am not surprised by Professor Terman's announcement (p. 56) that Binet made the same distinction fifteen years ago; but when Professor Terman goes on to say that "it has been especially insisted upon by all psychologists in the field of individual differences" I am mildly curious to know why the many competent psychologists present at the Chicago meeting did not recognize their old friend and use it then and there to demonstrate that the "bombshell" was only a "dud" after all. Prompt action at that time might have saved the expense to which one of the publishing houses specializing in test materials went in reprinting Professor Terman's editorial and sending it broadcast.

All this impels me to believe that the particular distinction that I had in mind has not been so clearly recognized as it might be—or, at least, that its implications have not been so thoroughly explored as the gravity of the issues at stake clearly demands. These implications seem to be tunnels into the very foundations to deterministic philosophy. Is it not indicative of its hectic haste to create out of hand a predictive science that this school of psychology has left the citadel of its central hypothesis so easily open to attack?

May I call attention to a few of the implications in question? Let us agree that what the determinists call "general intelligence" is either a single or multiple biological trait. If so, it is *a priori* improbable that its general level has been essentially changed during the course of human history, or, indeed, during the long prehistoric period that elapsed between the completed evolution of man's present structure and the invention of writing. There is, too, an abundance of empirical evidence to support this inference—for example, the superb cranial development that the Crô-Magnon race attained fully 20,000 years ago. If, as Professor Terman insists, "intellectual abilities are pretty largely determined by native endowment" (p. 60), it seems equally true that general intelligence took an unconscionably long time in getting into action, and the suspicion is aroused that present-day psychology is ascribing vastly more significance to nature and far less significance to nurture than the facts warrant.

This suspicion is strengthened by a mass of evidence so overwhelming in the aggregate that the suspicion rapidly grows into a conviction. If, for example, native intelligence were so superior a force as compared with culture, we should expect to find innumerable instances of discontinuity in social evolution; we should expect that superior intelligences widely separated in space and time would have "hit upon" the same improvements in the conditions of life. As a matter of fact, the evidence points compellingly toward a continuity of culture even in prehistoric times.[2] It has clearly been the accumulation and consolidation of experiential gains that has lifted humankind from the paleolithic levels.

Again, if native intelligence were so all-important, the turning-points of social evolution would have been marked by positive variations or increases in native intelligence and these would necessarily have been widely spread among millions of individuals. This hypothesis not only involves us at once in a biological absurdity, but is clearly inconsistent with the facts. The great turning-points of social evolution have actually been marked by improved methods of disseminating experience—*of letting more light into common minds!* The invention of writing has been characterized quite without exaggeration as the greatest event in human history. It is no mere coincidence that the beginning of "modern" history was contemporaneous with the development of the printing-press in Europe. Nor is it at all accidental that the nations today that rank highest on the scale of civilization and progress are those in which a substantial form of elementary *popular* education is most nearly universal.

I am aware, of course, that efforts have been made to explain social evolution in terms of biological improvement. It is easy to say that certain countries are backward because their peoples are either biologically undeveloped or biologically decadent. Galton explained the Dark Ages on the assumption that the high I.Q.'s of both sexes sought a celibate refuge in the monasteries and convents, leaving the common man, the moron, and the imbecile to propagate the species. Mr. Madison Grant and others have attributed all human progress to the native abilities (not exclusively or even predominantly intellectual, by the way) of the blue-eyed Nordic peoples, and have voiced a solemn warning that, unless the Nordic stock keeps its own blood pure and the "inferior" brunette peoples in a proper degree of subjugation, civilization will quickly and inevitably "go to pot." On the other hand, however, we have the now classic conclusions of Odin, Ward, and Cattell, emphasizing the outstanding significance of environmental factors; more recently Huntington[3] has made out a plausible case for the environmental influence of climate as a prime determinant of human achievement; developments in the field of tropical medicine seem to explain in the humbler but less fatalistic terms of the hookworm and the malaria-mosquito many defects that the determinist would ascribe to the inborn incapacity of the tropical peoples; while the possibilities of controlling some of the physiological conditions of mental development through endocrine treatment, if not as yet warranting extravagant hopes,[4] are at least sufficiently convincing to counteract the equally extravagant despair that determinism encourages.

---

[2] Cf. a summary of the evidence, with citations to the monographic literature in the first supplementary volume of the *Encyclopaedia Britannica*, v. xxx, 1922, "Anthropology," especially pp. 147 ff.

The current teachings of educational determinism are dangerous because they proceed with a dogmatic disregard of the possibilities of insuring progress through environmental agencies. This disregard is so studied, so pointed, as to brand the determinist as thoroughly prejudiced. Take, for example, the current deterministic interpretations of the results of the Army tests — interpretations which leave entirely out of account any possible influence of environmental factors. The fact that the median mental age of the enlisted men coincided almost exactly with the median age of leaving school is completely ignored. The higher mentality of those who remained in school is dogmatically explained as due to selection with no whit of credit given to the school.[5] Even the clear possibility that practice-effect might transfer from recent school experiences in taking examinations to the similar test situation is never mentioned, although every experiment on "transfer" has taught us that practice-effect is a factor that must be reckoned with. Personally I am astounded that competent psychologists and trained students of science should be guilty of such loose reasoning. Long overdue is the work that was reported only recently by Mr. Herbert B. Alexander[6], who correlated Army Alpha scores taken by states with the rank-order of the states in respect to educational facilities and other environmental advantages. How the determinists are going to square themselves with the public in the face of these correlations puzzles me; I think that it will puzzle them a bit, for Mr. Alexander's conclusion is that, "in so far as it applies to such large social groups as the American states, Army Alpha appears as a test of what *has* been learned rather than what *can* be learned."

---

[3] In his *Civilization and Climate* (1915) and other works.

[4] See the admirable symposium on the endocrines in the *Journal of the American Medical Association*, 79:89-109, July 8, 1922.

[5] Professor Terman's own arguments against accepting the results of the Army tests as indicative of average adult mental age contain no suggestion that the differences found may have been due to educational inequalities. The educational factor is mentioned but in a context that shows clearly that Professor Terman is thinking of the high school and college as selective agencies. (See Terman, L. M. "Mental growth and I.Q." *Journal of Educational Psychology*, 12:330-331, September, 1921.)

Why is it that a conclusion so obvious to anyone acquainted both with the theory of the tests and with the wide variations in American educational systems was not considered as at least a possibility before the fatalistic and fantastic interpretations were sent broadcast? Rightly interpreted, the results of the Army tests constitute a fairly clear vindication of the contentions of the environmentalists. How does it happen that quite the opposite implications have gone forth, if not with the sanction of, certainly without a disclaimer from, the responsible group? At the present juncture in social evolution it should be no light matter under the guise of science to promulgate an obviously one-sided and now apparently fallacious theory the net effect of which is to weaken public faith in universal education.

Clearly if the possibilities of insuring progress through environmental agencies had only an equivalence of plausibility as compared with Galtonian fatalism, every consideration of humanity would compel us to accept them as working hypotheses in education. As a matter of fact, the plausibility is not equivalent, it is distinctly preponderant. Professor Terman suggests that Galtonian fatalism "is entitled to the right of way as a guide to action until something more substantial than sentiment can be brought to bear against it." Whether the facts and reasoned inferences that I have adduced are more than "sentiment," the reader will be able to judge. Should he consider them sufficient only to afford a fairly even balance with the opposing facts and inferences, then I contend that sentiment should be frankly permitted to cast the deciding vote. Sentiment has not been an entirely negative force in this world of ours. It has accomplished some rather important things right in the face of the theory that Professor Terman is supporting. What, may I ask, would have been the effect of the anti-slavery agitation if the hypothesis of an unmodifiable "general intelligence" had been current at that time? What would be the case of the universal franchise×? Indeed, why not hereditary leadership and even the divine right of kings if only these doctrines could be tempered with a little Mendelism? We are wont to grumble about the slowness with which science develops. Personally I am devoutly grateful that "sentiment" had an opportunity to work a few miracles before modern psychology "discovered" these innate, radical, and unchangeable differences in native mentality.

---

[6] Alexander, H. B. "A comparison of the ranks of American states in Army Alpha and in social economic status," *School and Society*, 16:388, September 30, 1922.

But to get back to the distinction between vertical and horizontal growth, which holds, I believe, the key to this apparent antinomy between nature and nurture. Granting a biological basis for mentality, and granting that this basis is subject to the biological laws of variation and growth, it is my contention (1) that the innate variations diminish rapidly *in social significance* as we go up the scale from the idiot toward the genius, because (2) horizontal growth (stimulated by environmental forces) compensates in many important ways for the differences in vertical growth (due to native factors). The relatively "low visibility" of the differences between the common man and the genius as compared with the differences between the idiot and the common man, has not been mentioned by any of the determinists who have been so kind as to discuss my paper. It is, I think, a crucial issue.

× The day will come when a feebleminded man will not be permitted to cancel that of the most intelligent and best man in the state.

Professor Terman barely approaches this issue when he denies (p. 59) that there is any "improvement in the ability to perform more and more difficult types of intellectual operation" or in the "ability to deal with the abstract" between the ages of 21 and 31, and cites the results of the Army tests in support of this sweeping statement. Professor Terman, I am then safe in assuming, holds that ability to deal with the abstract increases up to the limits of vertical (biological) growth, and then suddenly stops. His evidence for this is drawn from group tests of a relatively simple character given only once to men between the ages in question. He implies that there was no correlation between age and intelligence as thus measured.

Professor Terman hopes that, if I believe this "negative result," to be due to the "use of a faulty instrument," I shall "devise another capable of yielding more valid data." I wish that I had both the time and the intelligence thus to help my good friends, the determinists, out of what seems to be a pretty bad hole. Somebody ought to do it.

Does it not seem strange, however, that Professor Terman lays so much stress on the results of the Army tests when they seem to uphold his preconceptions as to the absence of mental growth between 21 and 31, and yet vigorously discredits them when they go counter to his contention regarding the average mental age of adults?[7] In the latter controversy (as to whether mental maturity comes in the average of cases at 13.4 years or 14.5 years or 16 years), I am just now interested only as it typifies the fog in which the whole theory of the I.Q. is enveloped. Professor Terman's assertion that the Army tests do not show any evidence of intellectual growth between the ages of 21 and 31, however, is worthy of further attention. Reference to the bulky report of the Army tests discloses the fact[8] that no correlations were worked out between the scores made by the enlisted men and their chronological ages. In the case of the officers, the tabulations of scores according to age show a steady decline in "intelligence" between the ages of 20 and 60. According to the graph (Memoirs XV, page 814) the officers under 21 are the most intelligent of all. The officers between 51 and 60 show a deterioration of about twenty percent from the 20-year level.

---

[7] Terman, L. M. "Mental growth and the I.Q." *Journal of Educational Psychology,* 12:325, 401, September, October, 1921. (Especially pages 330-331.)

[8] Yerkes, Robert. (editor) *Psychological examining in the United States Army.* (Memoirs of National Academy of Science, Vol. 15) Washington, D. C.: Government Printing Office, 1921. p. 813.

The deterministic explanations of this apparently progressive deterioration of intelligence with age illustrate admirably the inevitable tendency of deterministic thinking. The obvious explanation (which the determinist studiously avoids) is that the Alpha tests are tests of education, and that, in consequence, the younger officers, being closer to their school days, do far better in the tests than do the older officers. It is possible, too, that the improved schooling of the past two decades has helped the younger officers. Even to suggest this, however, would be impossible to the dyed-in-the-wool determinist. All differences ***must*** be due to "selection" if the hereditarian hypothesis is to be upheld—consequently all other possibilities are rigorously excluded!

It is conceivable that the Army tests might have given us some valuable information concerning the possible growth in intelligence (or whatever it is that the Alpha tests measure) between the ages of 21 and 31. If the test scores of enlisted men coming from the clerical occupations which give constant practice in many of the Alpha materials were tabulated according to age, it is possible that a positive correlation might be found. Indeed, this is fairly predictable since the Alpha scores always "correlate" highly with length of schooling—a fact that the determinist again attributes not to an assumption that Alpha may measure education, but to the overworked hypothesis that (education always and inevitably selects the bright minds)

Let us drop the Army tests, however, and return to Professor Terman's clear implication that even the most specific training in abstract thinking will not improve "one's ability to perform more and more difficult types of intellectual problems" or one's ability to "deal with the abstract," once biological maturity has been reached. Lincoln at twenty-one, it would seem, had as much ability to grasp tremendously complicated and difficult problems as he had at fifty-two! Darwin on the Beagle at twenty-two had the same capacity for dealing with abstract biological questions that he possessed twenty-eight years later when he published the *Origin of Species!* All the intervening years of struggle, of study, of experience with men, added nothing to Lincoln's ability; all the intervening years of observation, experiment, and close thinking did not increase Darwin's acumen by so much as an iota!

Experience of effect? [illegible]

"Nonsense," the determinist will protest, "that is not what I mean. I only mean——." But I shall let him tell us what he means—if he can. I shall only say that both Darwin and Lincoln could do a type of thinking at fifty that they could not do at twenty-one—or at thirty-one, for that matter. I contend that the growth in the intervening years was real intellectual growth, and if it be objected that it was not growth in "general intelligence" I answer that it then represented something vastly more important than general intelligence.

I furthermore contend that intellectual growth of this horizontal type is possible among individuals of very moderate I.Q.'s. My contention is not merely an "emphatic assertion," as Professor Terman would have us believe. It is sanctioned not only by common-sense and everyday experience but by practically every recorded experiment in learning. That specific abilities are greatly increased by training I thought everyone conceded; the only question has seemed to be how far the improvement "transfers" to other situations. Even this is still an open question; but Professor Terman's sweeping statement is a clear denial of the potency of *specific* discipline itself, once the limits of vertical growth have been reached.

Professor Terman wonders whether my "apparent faith in pronouncements based upon inner vision" would let me give "much weight to any kind of objective evidence the experimentalist might be able to bring forward." I make no pretense of being a man of science; I confess, indeed, to abounding faith in certain "pronouncements based upon inner vision," and I have small faith in what the scientist calls the "naked fact," preferring those that are decently clothed with meaning. However, if my scientific training failed to make me a scientist, it left, I think, as a faint deposit, a moderate ability to distinguish fairly clearly between a fact and an artifact, between a fact and an assumption, and between a law and an hypothesis. Some frank untangling of their "evidence" on the basis of these distinctions, I venture to say, might well have been for the determinists a preliminary to the rather presumptuous assumption that, in the field of education, they and they alone are doing truly "scientific" work.

And speaking of facts and meanings leads me to another disagreement with Professor Terman. He does not like my "naïve" definition of democracy as the collective supremacy of the common man. Nor do some other people. I did not formulate this definition as an ideal, but as a fact—and a most tremendous fact fraught with a most tremendous meaning. Professor Terman implies that the collective supremacy of the common man is inconsistent with the facts of biology and psychology. If so, a big fact is in conflict with some little facts. As I have suggested, practically every great forward movement in social evolution has been made in the very face of the theory for which Professor Terman stands. Each of these movements has tended toward the liberation of the great masses and the assumption by them of larger and larger measures of power. After studying this development with apparent justice from every point of view, Lord Bryce concluded that democracy, even in this "naïve" sense, had abundantly justified itself. The big fact seems to have swallowed and digested the little facts, and theories based on the latter and utterly disregarding the former would consequently seem to have now only an academic interest. Professor Terman wishes to formulate a definition of democracy that will "square with the demonstrable facts of biological and psychological science." (p. 62.) I admire this ambition of the determinists to set an example of clear thinking. Is it unkind to suggest that their own field fairly bristles with opportunities for practice? Democracy, as the collective rule of the common man, seems likely to get on fairly well for a while on the present definition; but a good, sound, clear definition of "general intelligence" that will "square" with common-sense, everyday experience, and the facts of history would be a "winner."

Such a definition, if happily achieved, might do something to clarify the determinist's conception of "leadership" and the alleged responsibility of the school to select potential leaders and equip them for social service. Professor Terman says that I would consider it "most undemocratic to remind a class of college graduates that they are marked by ability and training for leadership." (p. 61.) Quite so; I would rather tell them that they owe to society a debt of *service* in return for the privileges and advantages that they have enjoyed. Leadership in any sense of the term they may or may not achieve, but positive service of some sort they can all render. It seems indeed to be the verdict of human experience that the less they have their eyes on the "distinctions" and "prominence" that the term "leadership" so inevitably implies, the greater will be their chance of becoming effective leaders in a democratic society. Nothing seems to be more inimical to such leadership than an over-weening consciousness of one's superiority to the common run of humanity. This negative and "inferior" trait, it seems, more than counterbalances the traits, however numerous, that are positive and constructive. Certainly a striking characteristic of the very leaders whom Professor Terman names is their humility—their sincere humility, not a posed substitute. In my judgment, neither social progress nor individual happiness will be augmented by encouraging young people to believe that they were born to lead. With individuals as with nations, the superiority-complex seems to be heavily charged with dynamite.

If it is desirable that there should be more high I.Q.'s among our effective leaders, the best way to get them, I contend, is to educate the rank and file justly to evaluate and select them rather than to imbue these gifted children with the notion that they were Heaven-sent to lead their dull fellows. Professor Terman accuses me of confusing "leadership" with "drivership." Well, whether the common man follows blindly or is blindly driven seems to me quite inconsequential. It is the tacit and nonchalant assumption of necessary and irremediable blindness that I am calling into question. It is my contention that the common mind of humanity has already demonstrated its ability to think for itself; that universal education can train it to think more clearly and in larger units—in any case, that universal education can give it a common stock of dependable ideas with which to do its collective thinking—and that the first and foremost task of education is to do this job passing well.

I should be remiss if I were to close this rejoinder without again stating my conviction that mental tests have made very important contributions to educational progress. It would be the height of folly for education not to avail itself of these instruments in detecting individual differences in learning capacity whether such differences be innate or acquired. My quarrel is not with the tests, but with the fatalistic assumptions which are part of their "heredity." They derive from Galton, and they are over-burdened with Galtonian tradition. The present-day determinists repudiate some of Galton's teachings. One in particular is to them anathema, although it seems to me far worthier of survival than others that they retain. I refer to Galton's theory that genius is eminently well qualified to look after itself—that it is bound to come into its kingdom in spite of almost every obstacle short of premature death. I have a suspicion that Galton came pretty close to the truth here. On the other hand, there is the great rank and file of common or average intelligence. To endow the masses with genius is biologically impossible; but to endow the masses with the fruits of genius is both educationally possible and socially most profitable. The mental tests will help most if they aid the teacher in discharging this transcendent duty. They will render a gratuitous and disastrous disservice if they encourage in the teacher the conviction that the illumination of common minds is either an impossible or a relatively unimportant task.

This is the "man of straw" against which (or whom?) Professor Terman accuses me of waging a "spectacular fight." (p. 60.) Perhaps the straw stuffing got damp and fermented—in any case, I notice that it has developed a very lively "kick." And is it not passing strange that three important programs at the Chicago meeting should have been devoted to a series of attacks upon a straw man, and that these attacks were directed chiefly by specialists in the field of mental measurements? It should be "read into the record" that at least two of these programs were either instigated or deliberately planned by the psychologists to offset what they themselves believed to be dangerous tendencies in the use of the tests and the interpretations of the results. My own criticisms of the movement were mild as compared with the appalling fallacies and inconsistencies pointed out by recognized authorities in the field itself, and especially by Messrs. Colvin, Rugg, and Trabue. As an "outsider," I have been made the "goat"—but to be even a goat in a good cause is better than to be a sheep blindly following the bell-wether in a bad cause that happens at the moment to be the fashion.

PROGNOSIS TESTS OF ABILITY TO LEARN FOREIGN LANGUAGES

THOMAS H. BRIGGS
Teachers College, Columbia University

Among school administrators there is a widespread opinion that too large a proportion of secondary-school pupils are permitted to elect a foreign language. Certainly the numbers who drop the subject without continuing it long enough to achieve any permanent power manifest one element of waste that should so far as possible be lessened. The increase of junior high schools with their earlier offerings of foreign languages serves to exaggerate the seriousness of the situation. The two chief possibilities of economy are (1) the modification of foreign language courses so that they may be of value to the extent to which pursued and (2) the restriction of election in this field to those who are known to have the ability to profit by it.

In 1918 Professor Truman Lee Kelley and I attempted to ascertain the extent to which a prognosis could be made by tests of the ability to learn foreign languages. For good reasons it was not possible to prosecute the study far enough to make the results of great value for practical use, but it is thought that the method used and the tentative conclusions reached warrant publication at this time. It is hoped to continue the study later in application to each of the foreign languages offered generally in our secondary schools.

The first step was the analysis of learning a foreign language into the major mental activities involved. A pupil who studies any foreign language in our schools is required (1) to memorize a vocabulary, (2) to memorize paradigms, and (3) to translate. Other activities are also required, but these three were selected as being the most important in the eyes of the teachers, who judge of failure or success.

The second step was to find or to devise tests that seemed likely to measure the abilities involved in the major phases of the work. The tests selected, with a brief justification of their choice, are as follows:

1. *Free association.* The pupils were required to write in columns as many different words as they could in three minutes. The total score was the number of words written. This test was used chiefly on the assumption that the same mental powers that had operated in the acquisition and recall of English words would function in learning and using the vocabulary of a foreign language.

2. *Extent of vocabulary.* The pupils were required to mark after each of 100 words, which had been selected at regular intervals from a standard dictionary, a "1" if they knew the meaning. At the conclusion of these responses the pupils were asked to write a definition of each of twenty designated words that they had marked "1." The percent of correct definitions was used as a correction index for the total score on the original 100 words. This test was used on the same assumption as that underlying the free association test.

3. *Memory of nonsense syllables.* An original test devised by Professor Kelley. It was used on the assumption that at least in the beginning of the study of a foreign language the acquisition of vocabulary and paradigms is similar to the rote memory of nonsense syllables.

4. *Briggs Analogies.* This test required the pupils to write a fourth word which bears the same relation to a third that the second does to the first. All the relations are of the forms of nouns, pronouns, and verbs, e.g.,

boy	boys	woman	....
cow	cow's	she	....
give	given	sit	....

This test was used chiefly on the assumption that the same mental powers involved in learning the forms of English words would be used in learning the forms of foreign words. Moreover, used in other experiments, the test had manifested a high reliability and close correlation with measures of general intelligence.

5. *Kelley-Trabue Completion Test.* This well-known test was used chiefly on the assumption that it measures powers that are involved in translation, that to insert "wet" in "The boy in the rain got ...." is similar to supplying the meaning in a foreign language sentence, the meaning of which is partly known, to selecting among several possible meanings; or to deciding whether *horti* is a genitive singular or a nominative plural. Moreover, the test has a high reliability and has correlated well with measures of intelligence.

6. *Kelley Substitution.* This test, devised especially for the study, requires the substitution of certain symbols for stimulus numbers. The faster and more accurately the key at the top of the page is learned, the higher the possible score. This activity was thought to be similar to translating, especially from English into a foreign language.

7. *Kelley Opposites.* An opposites test was used because of the fact that it is wholly linguistic and that in another study it had given a good positive correlation with pupils' general success in school work. For this study Kelley devised a new test of forty paired words, graded in difficulty, the pupils being required to indicate whether the meanings were the same or the opposite.

This battery of seven tests was administered to a number of entering pupils who elected a foreign language in seven high schools in Parkersburg, West Virginia; Topeka, Kansas; New Brunswick, New Jersey; Meriden, Connecticut; Mankato, Minnesota; Cleveland, Ohio; and New York City. The tests were given and scored by the teachers.

The only criterion of success used was the term marks in foreign language. These marks are so unreliable that the resulting coefficients of correlation can be considered only as "promising" and indicative of the probable fact that the method would produce much better results with a more stable criterion. To attempt to correlate with teachers' marks is like an attempt of a blind man to walk sociably with a drunkard over very uneven ground without physical contact. Kelley has reported[1] that the correlation between the marks of two teachers for the same pupils in the same subjects ranges from 0.28 to 0.38, and he computed that the highest correlation that might be expected from prognosis tests with teachers' marks would be the square root of the reliability of the latter. In so far as his earlier estimate of the reliability of teachers' marks is true for the teachers who cooperated in this study, the theoretical limit of the possible correlation between the prognosis tests and foreign language marks is 0.53 to 0.62.

There was no means of knowing whether the cooperating teachers were more or less skilled and accurate than the average in estimating pupils' success. We do know that the results of the several tests correlate very differently with the marks assigned by the individual teachers. The scores of the analogies test, for example, correlated 0.025, 0.307, and 0.454, and of the completion test 0.142, 0.379, and 0.796 with the term marks assigned by three teachers in the same school.

By means of a regression equation, Kelley selected a battery of three tests which when weighted as follows gave the best results:

Briggs Analogies (4)
Kelley-Trabue Completion (2)
Free Association (1)

The correlation of the scores on these three tests with the marks of all teachers for the 598 pupils in the six high schools outside New York City who remained until the end of the semester was 0.47. As evidence that the tests were measuring special ability to learn foreign languages, the weighted scores of the battery were correlated with pupils' marks in English. The correlation was 0.31—distinctly lower.

Considering the theoretical limits previously stated, a correlation of 0.47 is quite as high as one could hope to get; but is far too low to use alone as an administrative device for excluding those who have small probability of succeeding in the study of foreign languages. The battery might be used profitably and practically, however, in connection with other measures of pupils' probable expectation of remaining in school, etc. Table I shows that of the pupils having the highest quarter of the battery scores, 71.1 percent received marks of 80 or better; while only 18.1 percent of those in the lowest quarter received similar high marks. One-third of those in the lowest quarter failed, as compared with only 3.4 percent of those in the highest. Of the 299 pupils who were in the upper half of the prognosis test scores, only 12 failed; but of the 150 in the lowest quarter two-thirds passed in the first term's work.

---

[1] Kelley, T. L. *Educational guidance.* (Teachers College Contributions to Education No. 71) New York: Teachers College, Columbia University, 1914. p. 15.

TABLE I. PERCENTS OF THE PUPILS IN EACH QUARTER OF THE PROGNOSIS TEST SCORES WHO RECEIVED THE INDICATED SCHOOL MARKS IN FOREIGN LANGUAGES

(598 *Pupils in Six High Schools*)

School Marks	Prognosis Test Scores			
	First quarter (lowest)	Second quarter	Third quarter	Fourth quarter (highest)
Failed (0–69)	33.6	20.8	4.8	3.4
Passing (70–79)	48.3	44.1	41.8	25.5
Superior (80–100)	18.1	35.1	53.4	71.1
	100.0	100.0	100.0	100.0
Highly Superior (90–100)	3.4	5.2	11.6	30.2

Table II shows the percents of the failing, passing, and superior pupils who fell into each of the four quarters of the prognosis test scores. With the exception of the scores in the third quarter, they fall into regular straight line curves.

The results with the 304 pupils in the High School of Commerce of New York City were not so satisfactory, though they show that the battery of tests does to some extent measure the

TABLE II. PERCENTS OF THE PUPILS IN EACH SCHOLARSHIP GROUP WHO FELL INTO THE SEVERAL QUARTERS OF THE PROGNOSIS TEST SCORES

(598 *Pupils in Six High Schools*)

Prognosis Test Scores	School Marks			
	0–69	70–79	80–100	90–100
Lowest quarter	53.2	30.3	10.2	6.7
Second quarter	34.0	26.6	20.4	10.7
Third quarter	7.4	27.0	29.4	22.7
Highest quarter	5.3	16.1	40.0	60.0
	99.9	100.0	100.0	100.1

ability of ninth-grade pupils to learn a foreign language. Unfortunately the detailed data have been lost so that the coefficient of correlation cannot be given; but the collected grouping of the pupils by quartiles is shown in Table III. The passing mark in this school is 60, instead of 70, as in the others. Why the results were less satisfactory in this school than in the other six one may not with any certainty say. There are factors, however,

TABLE III. PERCENTS OF THE PUPILS IN EACH QUARTER OF THE PROGNOSIS TEST SCORES WHO RECEIVED THE INDICATED SCHOOL MARKS IN FOREIGN LANGUAGES

(304 *Pupils in High School of Commerce, New York City*)

School Marks	Prognosis Test Scores			
	First quarter (lowest)	Second quarter	Third quarter	Fourth quarter (highest)
Failed (0–59)	23.6	30.3	29.8	10.1
Passing (60–79)	58.3	47.4	44.8	36.2
Superior (80–100)	18.1	22.4	25.2	53.7
	100.0	100.1	99.8	100.0
Highly Superior (90–100)	2.8	6.6	8.0	24.7

which make it probable that the battery should have been no less effective in prophesying the ability of these metropolitan pupils to acquire the rudiments of foreign languages.

Certain good reasons made it impracticable to apply the battery of three tests to other pupils or to refine and supplement the method used. It is desirable (1) to divide the selected tests into equal halves, each administered at a different time, so as to obtain an estimate of the effect of accidental factors; (2) to consider the scores and the class marks of pupils who, for one reason or another, are eliminated before the end of the semester; (3) to make further correlations of the test scores with marks at the end of the second semester and between the marks of the first semester and those of the second; (4) to gather information concerning all the pupils as to absence, illness, outside work, any foreign language spoken or heard at home, effort, etc., with "case studies" of those whose marks are at variance with their test scores; and (5) to prepare a set of criterion tests more reliable than teachers' marks and more indicative of the knowledge acquired in foreign languages. Most of these improvements have been used by one of my students, Dr. William Sims Allen, in a study soon to be published on the prognosis of ability in first-year Latin.

COMPARATIVE VALIDITY OF THE HOTZ SCALES AND THE RUGG-CLARK TESTS IN ALGEBRA

ELEANORA HARRIS
Central Missouri State Teachers College

FREDERICK S. BREED
University of Chicago

In his introduction of "A Critical Study of the Harvard-Newton Composition Scales," Isidore Kayfetz says: "In this country many professors, students, superintendents, principals, and teachers are at present engaged in writing books, monographs or articles on standard tests and scales. All who are so engaged may be roughly divided into three classes: (1) those who devise scales, (2) those who praise them, (3) those who condemn them. There is a crying need for a fourth class, namely, those who will study them historically, critically, comparatively, and experimentally in the light of experimental-pedagogical criteria and methods. Perhaps the future will give rise to this class." As members of this fourth class, the writers undertook this investigation.

Three of the First Year Algebra Scales, Series A, of H. G. Hotz and seven of the Standardized Tests in First-Year Algebra of H. O. Rugg and J. R. Clark were selected for study. The Hotz scales are development tests. The Rugg-Clark tests, with the exception of the Verbal Problems, are rate tests.

The purposes of the study were: (1) to investigate the validity of both series of tests; (2) to compare the two series in regard to economy of time; (3) to study the difficulties in taking, giving, and scoring; and (4) to determine the value of the tests in diagnosis.

The investigation was made possible by the helpful co-operation of the principals, teachers, and pupils in the following Illinois high schools: Cicero, Elgin, Joliet, Kenilworth-New Trier Township, Maywood-Proviso Township, Whiting; and (at Chicago) the Harrison Technical, Hyde Park, Lane Technical, Parker, Wendell Phillips, and the University High School of the University of Chicago.

METHOD AND MATERIAL

The Hotz Problem Scale and the Rugg-Clark problem test, called Verbal List A, were given in ten classes. These classes, I-X, completed first-year algebra at the end of the first half-year. The tests were given the latter part of the following March and the first half of April. The time allowed for each test was 25 minutes. Acceptable papers were received in each test from 240 pupils.

In twenty-five classes were given the Hotz Equation and Formula Scale and the following seven of the Rugg-Clark Formal Tests: Preliminary Tests, A and B, and Tests 4 (Simple Equations), 10 (Fractional Equations), 11 (Formulae), 12 (Quadratic Equations), 13 (Simultaneous Equations). Eleven of these classes, XI-XXI, completed first-year algebra at the end of the first half-year. In these classes the compared tests were given the latter part of March and the first half of April. Acceptable papers were received in each test from 264 pupils. The same tests were also given in May of the same year in fourteen classes, XXII-XXXV, in which first-year algebra was completed at the end of the school year. Acceptable papers were received in each test from 267 pupils in these fourteen classes. Thus for the equation and formula tests, given in the twenty-five classes, XI-XXXV, 531 acceptable papers were obtained in both the Hotz and Rugg-Clark tests.

The Hotz Graph Scale and the Rugg-Clark Graph Test were given in eleven classes, XXXVI-XLVI. These classes completed first-year algebra at the end of the first half-year. The tests were given the last two weeks in March and the first part of April. Acceptable papers were received in each test from 283 pupils.

The tests of the different authors were administered in every class on the same day or on successive days. In twenty-three classes the Hotz test was given first; and in twenty-three classes the Rugg-Clark test was given first. Usable papers from tests by different authors were thus obtained from each of the same 1054 pupils.

Copies of the Rugg-Clark problem test, Verbal List A, were specially prepared for use in the investigation. Mr. Clark kindly gave instructions as to spacing and page arrangement. Copies of all the other tests were used in published form. In the administration and scoring of tests, the directions provided by the authors were followed carefully. All the tests were administered and scored by Miss Harris.

PROBLEM TESTS

Order of difficulty of the problems.—Tables I and II and the accompanying Figure 1 indicate the order of difficulty of the

TABLE I. HOTZ PROBLEM SCALE. PERCENT OF PUPILS SOLVING EACH PROBLEM

Problem Number	Hotz given first (Classes I-V)	Rugg-Clark given first (Classes VI-X)	Total (Classes I-X)
1	84.5	98.4	91.7
2	75.9	90.3	83.3
4	74.1	84.7	79.6
5	56.9	68.5	62.9
7	25.0	36.3	30.8
8	19.8	24.2	22.1
9	8.6	12.1	10.4
12	1.7	3.2	2.5
13	0.9	1.6	1.3
14	0.9	3.2	2.1

TABLE II. RUGG-CLARK PROBLEM TEST. PERCENT OF PUPILS SOLVING EACH PROBLEM

Problem Number	Hotz given first (Classes I-V)	Rugg-Clark given first (Classes VI-X)	Total (Classes I-X)
1	79.3	79.0	79.2
2	71.6	79.0	75.4
3	70.7	66.1	68.3
4	53.4	70.2	62.9
5	38.8	50.4	44.6
6	35.3	43.5	39.6
7	25.0	18.5	21.7
8	31.1	23.4	27.5
9	5.2	5.6	5.4
10	11.9	23.0	17.7
11	14.5	31.1	23.1
12	0.0	0.8	0.4
13	1.7	2.0	1.9

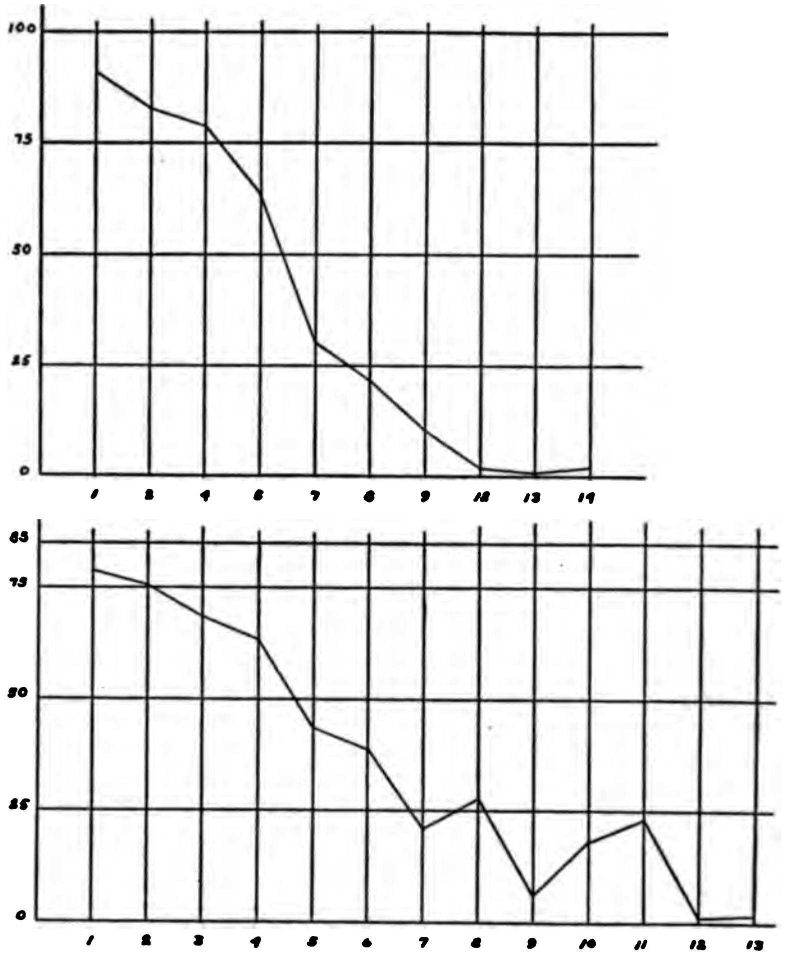


FIGURE 1. PERCENT OF PUPILS SOLVING EACH PROBLEM IN THE PROBLEM TESTS. UPPER CURVE FOR THE HOTZ TEST; LOWER CURVE FOR THE RUGG-CLARK TEST. DATA FROM TABLES I AND II

problems in the Hotz Problem Scale and the Rugg-Clark Verbal List A for the 240 pupils taking these tests. They show a wider range in the percents of pupils solving each problem of the Hotz test than for the Rugg-Clark test. The range for the Hotz test is 90.4 and for the Rugg-Clark test 78.8.

In the Hotz Scale the order of arrangement was found to be one of increasing difficulty, except that problem 14 was somewhat easier than problem 13. From problem 8 to problem 13 there was much irregularity of arrangement in the Rugg-Clark test.

For ten classes and two methods of scoring the average coefficient of correlation between the two problem tests was 0.61.

Distribution of achievement of pupils.—Tables III and IV and the accompanying Figure 2 show the distribution of achievement of the 240 pupils in the solution of problem in the Hotz Problem Scale and the Rugg-Clark Verbal List A.

Table III shows that 2.6 percent of the pupils in classes I-V failed to solve a single problem in this test. Every pupil in classes VI-X solved one or more problems. Table III also gives the total number of problems solved, the number of pupils, the arithmetic mean, the median score, and the author's tentative median standards of achievement.

The method of scoring the papers in each test must be kept in mind in comparing the distribution of pupils around the median in number of problems right as shown in Tables III and IV

TABLE III. HOTZ PROBLEM SCALE. PERCENTAGE DISTRIBUTION OF PUPILS ACCORDING TO THE NUMBER OF PROBLEMS SOLVED

Number of Problems Solved Correctly	Hotz given first (Classes I-V)	Rugg-Clark given first (Classes VI-X)	Total (Classes I-X)
0	2.6	0.0	1.3
1	5.1	1.6	3.3
2	23.3	7.3	15.0
3	24.1	21.0	22.5
4	17.2	29.8	23.8
5	16.4	25.8	21.3
6	7.8	8.9	8.3
7	2.6	4.0	3.3
8	0.0	0.8	0.4
9	0.0	0.8	0.4
10	0.9	0.0	0.4
Number of "Rights"...	404.	524.	928.
Number of Pupils.....	116.	124.	240.
Arithmetic Mean......	3.48	4.23	3.87
Median..............	3.8	4.7	4.3
	3-Months	6-Months	9-Months
Tentative Median Standards.........	4.3	4.9	5.6

TABLE IV. RUGG-CLARK PROBLEM TEST. PERCENTAGE DISTRIBUTION OF PUPILS ACCORDING TO THE NUMBER OF PROBLEMS SOLVED

Number of Problems Solved Correctly	Hotz given first (Classes I-V)	Rugg-Clark given first (Classes VI-X)	Total (Classes I-X)
0	2.6	1.6	2.1
0.5	0.8	0.0	0.4
1	4.3	7.3	5.8
1.5	2.6	0.8	1.7
2	16.4	8.1	12.1
3	15.5	11.3	13.3
3.5	0.0	1.6	0.8
4	11.2	14.5	12.9
4.5	0.9	3.2	2.1
5	16.4	10.5	13.3
5.5	0.9	0.0	0.4
6	5.1	10.5	7.9
6.5	2.6	5.6	4.2
7	6.0	7.3	6.7
7.5	1.7	2.4	2.1
8	6.9	8.9	7.9
9	2.6	1.6	2.1
9.5	0.9	0.8	0.8
10	1.7	2.4	2.1
11	0.9	1.6	1.3
Number of "Rights"...	510.	611.	1121.
Number of Pupils.....	116.	124.	240.
Arithmetic Mean......	4.40	4.92	4.67
Approximate Median...	4.	5.	4.5
Computed Median.....	4.3	5.1	4.7

and Figure 2. In all the Rugg-Clark tests, if two answers were required in a problem and one answer was correct, one-half credit was given. This was done upon the advice of both authors, who were consulted personally. The unit on the horizontal axis of the Rugg-Clark curve is one-half problem or 0.5. In accordance with the author's directions, no partial credits were given in any of the Hotz Scales. The unit on the horizontal axis of the Hotz curve, Figure 2, is one problem or 1.

Figure 2 shows that the distribution of achievement for the Hotz Problem Scale was quite normal, while that for the Rugg-Clark test was quite irregular.

EQUATION AND FORMULA TESTS

Since the design of the Hotz Equation and Formula Scale and that of the Rugg-Clark Formal Tests were wholly different, the results will be taken up separately.

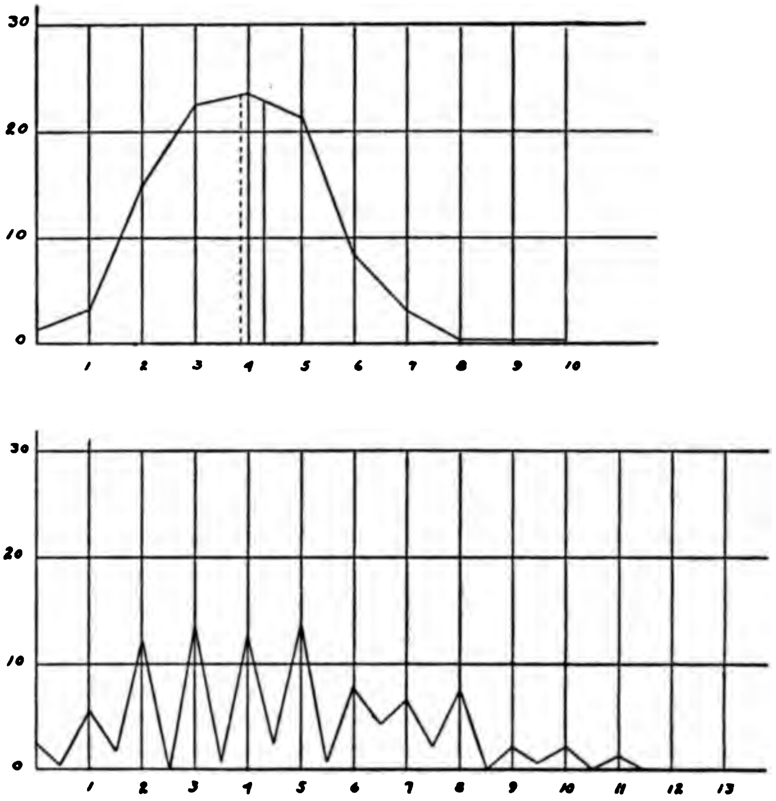


FIGURE 2. DISTRIBUTION OF PUPILS ACCORDING TO THE NUMBER OF PROBLEMS SOLVED CORRECTLY ON THE PROBLEM TESTS. UPPER CURVE FOR THE HOTZ TEST; LOWER CURVE FOR THE RUGG-CLARK TEST. DATA FROM TABLES III AND IV

Order of difficulty of problems in the Hotz Equation and Formula Scale.—Table V and Figure 3 indicate the order of difficulty of the problems in the Hotz scale for the 531 pupils taking this test. It will be recalled that the first group (classes XI-XXI) was tested some weeks after the work in first-year algebra had been completed and the second group (classes XXII-XXXV) a few weeks before the completion of it. It is very interesting to note the remarkable similarity of the curves representing the two groups. With the exception of problem 8, the

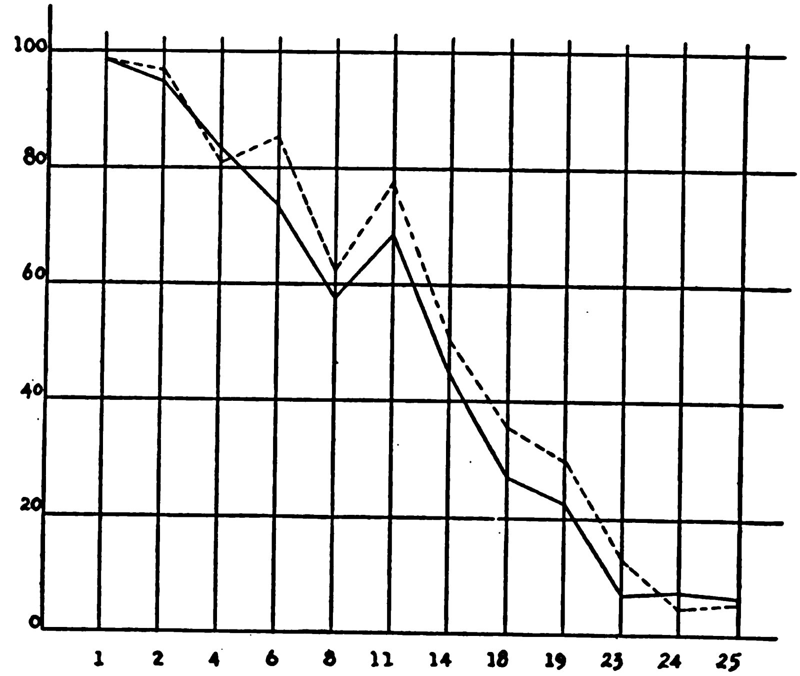


FIGURE 3. PERCENT OF PUPILS IN CLASSES XI-XXI (FULL-DRAWN LINE) AND IN CLASSES XXII-XXXV (BROKEN LINE) SOLVING EACH PROBLEM IN THE HOTZ EQUATION AND FORMULA TEST. DATA FROM TABLE V

TABLE V. HOTZ EQUATION AND FORMULA SCALE. PERCENT OF PUPILS SOLVING EACH PROBLEM

Problem Number	Classes XI-XXI	Classes XXII-XXXV
1	98.5	98.5
2	95.1	96.6
4	83.7	80.2
6	73.9	85.8
8	57.6	62.2
11	68.2	77.5
14	44.3	50.2
18	27.7	35.5
19	22.6	30.0
23	6.8	13.9
24	7.2	4.1
25	6.1	5.6

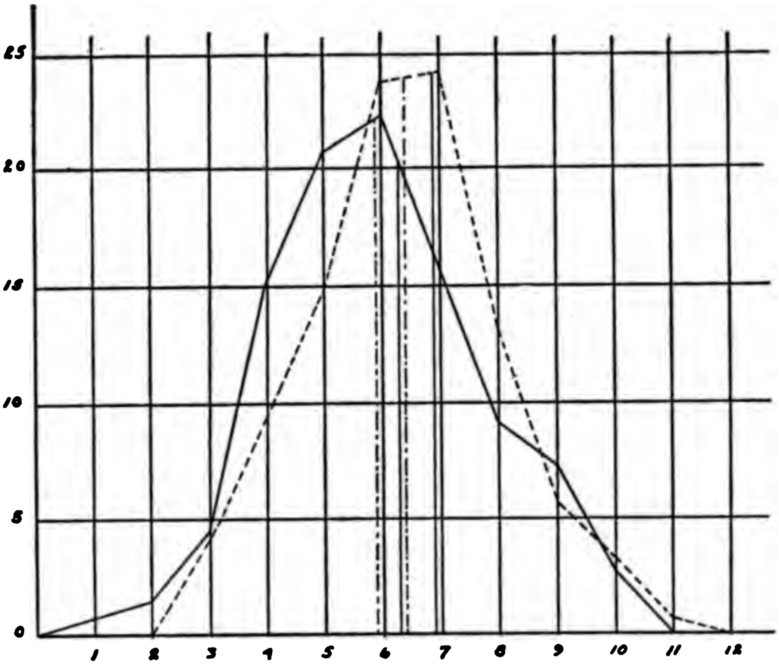


FIGURE 4. DISTRIBUTION OF PUPILS IN CLASSES XI-XXI (FULL-DRAWN LINE) AND IN CLASSES XXII-XXXV (BROKEN LINE) ACCORDING TO THE NUMBER OF PROBLEMS SOLVED IN THE HOTZ EQUATION AND FORMULA TEST. FULL-DRAWN VERTICALS SHOW MEDIANS AND BROKEN VERTICALS, ARITHMETIC MEANS. DATA FROM TABLE VI

fifth problem of the scale, the order of difficulty for these pupils corresponds to the order determined by the author.

Distribution of achievement of pupils in the Hotz Equation and Formula Scale.—Table VI and the accompanying Figure 4 show the distribution about the median for classes XI-XXI and XXII-XXXV. Of the pupils in classes XI-XXI, 0.8 percent succeeded in solving only one problem, while no pupil in any of the fourteen classes, XXII-XXXV, failed to solve at least three problems. No pupil in either group made a perfect score (12 problems right). Figure 4 shows that there is a remarkable similarity in the dispersion of the two groups and that they

TABLE VI. HOTZ EQUATION AND FORMULA SCALE. PERCENTAGE DISTRIBUTION OF PUPILS ACCORDING TO THE NUMBER OF PROBLEMS SOLVED

Number of Problems Solved Correctly	Classes XI-XXI	Classes XXII-XXXV
1	0.8	0.0
2	1.5	0.0
3	4.5	4.1
4	15.2	9.4
5	20.8	15.0
6	22.4	23.9
7	15.9	24.4
8	9.1	13.1
9	7.2	5.9
10	2.6	3.4
11	0.0	0.8
Number of "Rights"	1562.	1709.
Number of Pupils...	264.	267.
Arithmetic Mean...	5.92	6.40
Median..........	6.3	6.9

Tentative	3-months	6-months	9-months
Standards......	4.9	7.1	7.8

are comparable from the standpoint of distribution around the median.

Organization of cycles and evaluation of problems in the Rugg-Clark Formal Tests.—The results obtained from the twenty-five classes, XI-XXI and XXII-XXXV, when given the Rugg-Clark Formal Tests, may now be considered. In the design of these tests the authors used the "cycle" principle. The percent of pupils solving each problem in each cycle of each of the five tests was computed. The percents for Test 4 are shown in Table VII and in Figure 5. It is interesting to note how nearly equal the percents are for the two groups. The same thing was observed in the case of the Hotz Equation and Formula Scale. The number of problems solved of each type of problem in the cycles was also found. Percents were calculated showing what proportion the number of each type solved was of the whole number solved. These percents for Test 4 are given in Table VIII. A careful study of the percents, which are almost identical for the two groups, reveals the fact that the different types of problems in each test are of unequal difficulty for the

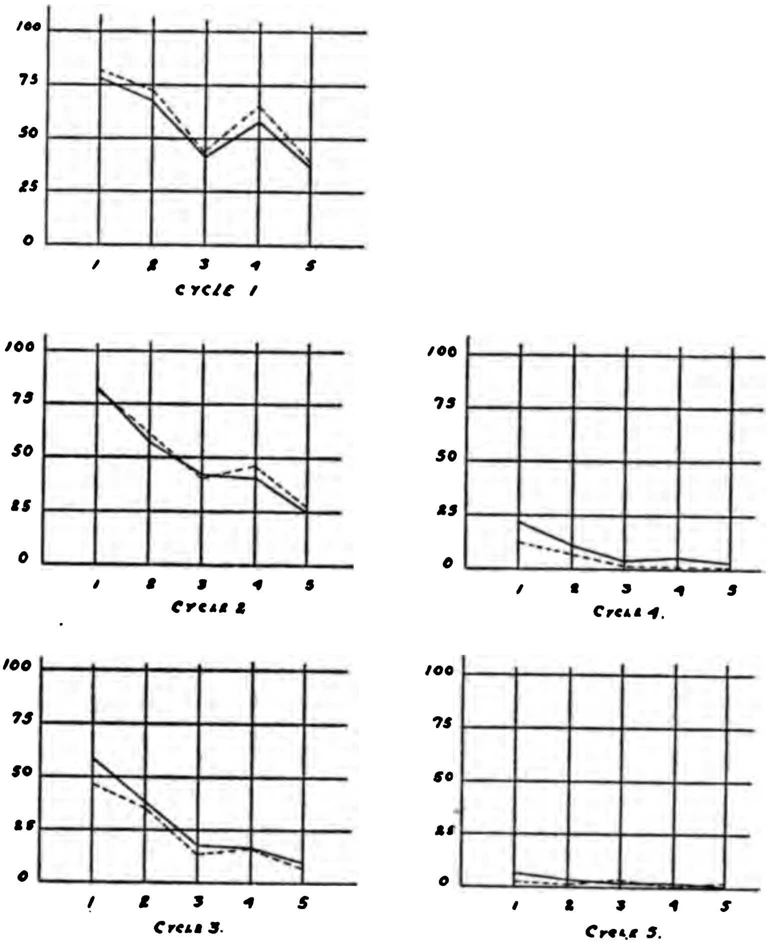


FIGURE 5. PERCENT OF PUPILS SOLVING CORRESPONDING PROBLEMS IN SUCCESSIVE CYCLES OF THE RUGG-CLARK TEST NO. 4 (SIMPLE EQUATIONS). DATA FROM TABLE VII

pupils tested. This is worthy of note in computing scores made by pupils taking these tests. Test 4 was selected because, of the five Formal Tests used, it is the only one designed in exact accordance with the "cycle" principle as described by the authors. Tables VII and VIII show clearly the relation existing between the total number of problems solved in this test and the number of each type of problem solved.

Variation in the rank of same classes measured by different instruments.—A study was also made of the variation in the rank of classes according to the tests of the different authors. Variations in both groups of classes were examined. Six of the eleven classes in the first group and five of the fourteen in the second either maintained the same rank or were displaced not more than one place. Some classes were displaced as much as five and six places. One class in the second group varied in rank 9.5 places. It is therefore not certain that the ability to solve algebraic equations is satisfactorily measured by either of the two tests.

TABLE VII. RUGG-CLARK FORMAL TEST 4 (SIMPLE EQUATIONS). PERCENTS OF PUPILS SOLVING CORRESPONDING PROBLEMS IN SUCCESSIVE CYCLES

Classes XI-XXI

Cycle Number	Problem in Each Cycle:				
	First	Second	Third	Fourth	Fifth
1	78.9	69.7	42.0	58.7	35.2
2	84.8	57.9	41.3	40.6	25.0
3	58.3	39.5	15.9	17.8	9.4
4	23.1	11.4	4.1	5.7	1.1
5	6.8	3.4	.8	1.9	....
Classes XXII-XXXV					
1	83.2	73.4	44.6	66.3	37.0
2	84.2	61.1	40.4	45.7	25.0
3	46.1	35.2	14.9	17.1	8.2
4	13.5	8.9	2.6	2.2	0.7
5	2.6	1.9	1.1	0.7	0.7

TABLE VIII. RUGG-CLARK FORMAL TESTS. PERCENT OF ENTIRE NUMBER OF CORRECT SOLUTIONS MADE ON EACH TYPE OF PROBLEM IN TEST NO. 4

Type of Problem	Classes XI-XXI	Classes XXII-XXXV
1	34.3	32.0
2	24.8	25.1
3	14.2	14.4
4	17.0	18.4
5	9.7	10.1

GRAPH TESTS

Order of difficulty of problems.—Table IX gives the percent of pupils in classes XXXVI to XLVI who solved each problem in the Hotz Graph Scale and the Rugg-Clark Graph Test. Figure 6 shows graphically the percents for the 283 pupils in each test. Five of the classes took the Hotz test first and six took the Rugg-Clark test first. In every class the two tests were given during the same period. The attitude of the pupils and a study of the papers received indicated that ample time was given for the Hotz test. It seems probable, therefore, that the percents given for this test indicate the order of difficulty of the problems for the pupils taking the test in this investigation. According to the results from the eleven classes and 283 pupils, the Hotz order of difficulty is affirmed in every case but one. In the Rugg-Clark test most of the pupils seemed helpless and were unable to use all of the time allotment. The results, however, substantiate the order of difficulty of the problems as reported by Rugg and Clark.

TABLE IX. PERCENT OF PUPILS SOLVING EACH PROBLEM. CLASSES XXXVI TO XLVI

HOTZ GRAPH SCALE		RUGG-CLARK GRAPH TEST	
Problem Number	Percent Solving	Problem Number	Percent Solving
1	94.7	1	10.8
3	75.3	2	4.9
4	86.6	3	4.2
5	51.9	4	2.5
6	47.3	5	0.9
7	36.4	6	0.2
10	15.9		
11	11.3		

Distribution of achievement of pupils.—Table X and Figure 7 show the distribution of achievement of the 283 pupils in the solution of problems in the two graph tests. The median number of problems solved in the Hotz test was 4.7. The median for the most efficient class, XLI, was 6.6. The author's tentative median standard is 5.6 for pupils who have studied algebra nine

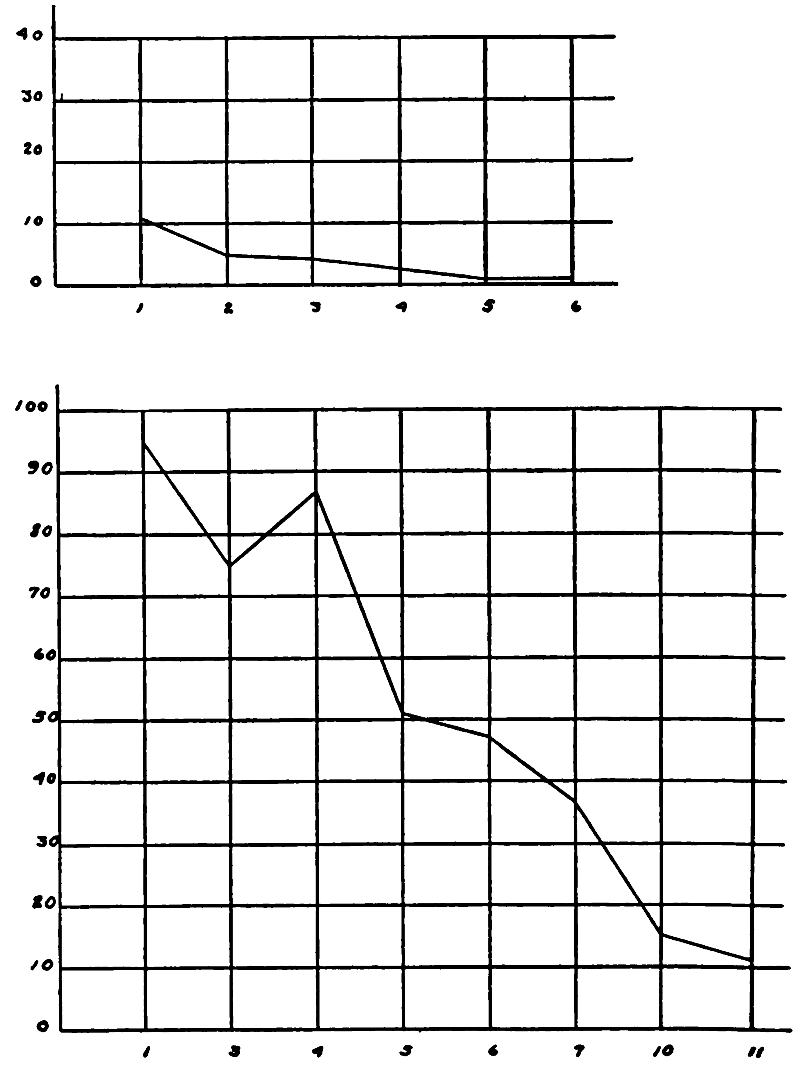


FIGURE 6. PERCENT OF PUPILS SOLVING EACH PROBLEM IN THE GRAPH TESTS. UPPER CURVE FOR THE RUGG-CLARK TEST; LOWER CURVE FOR THE HOTZ TEST. DATA FROM TABLE IX

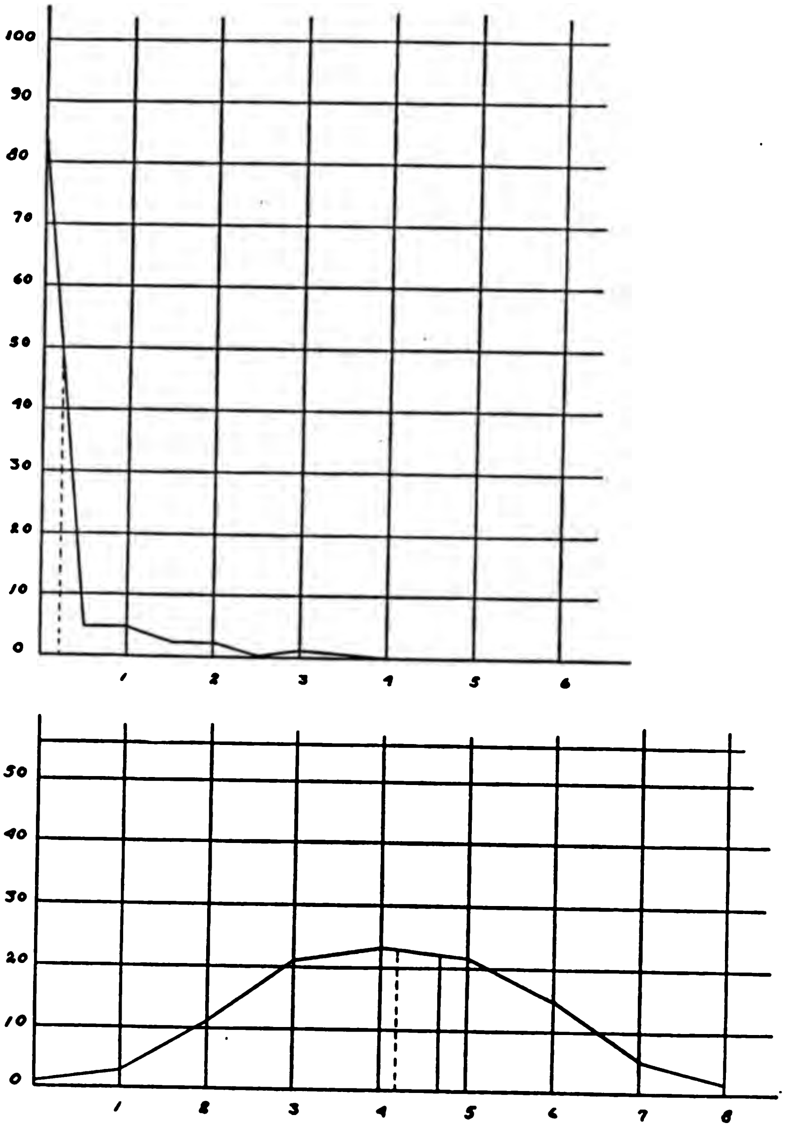


FIGURE 7. DISTRIBUTION OF PUPILS ACCORDING TO THE NUMBER OF PROBLEMS SOLVED IN THE GRAPH TEST. UPPER CURVE IS FOR THE RUGG-CLARK TEST; LOWER CURVE FOR THE HOTZ TEST. FULL-DRAWN LINES SHOW MEDIANS AND BROKEN LINES, ARITHMETIC MEANS. DATA FROM TABLE X

months. In the Rugg-Clark test the median score for class XLI was 1 problem solved in the 12 minutes allowed. The median score made in this test by all other classes was 0. The Rugg-Clark tentative standard, as determined by the average of 27 schools, is 1 problem for each 3.1 minutes. The 24 pupils in class XLI solved 30.5 of the 66.5 problems solved correctly by the 283 pupils in the Rugg-Clark test.

Tables IX and X and Figures 6 and 7 show very clearly that the Rugg-Clark Graph Test did not satisfactorily measure the pupils tested and that the Hotz Graph Scale did measure them. In the Rugg-Clark test 83 percent of the pupils failed to solve a single problem correctly. The above tables and figures show that the Hotz test gave a fairly normal distribution. Only 0.7 percent of the pupils failed to solve at least one problem and only 1.1 percent succeeded in making a perfect score.

TABLE X. PERCENTAGE DISTRIBUTION OF PUPILS ACCORDING TO THE NUMBER OF PROBLEMS SOLVED

Hotz Graph Scale		Rugg-Clark Graph Test	
Problems Solved	Percent	Problems Solved	Percent
0	0.7	0	83.0
1	2.5	0.5	5.3
2	11.0	1	4.6
3	20.1	1.5	2.1
4	23.0	2	2.1
5	21.9	2.5	0.7
6	14.1	3	1.1
7	5.6	3.5	0.7
8	1.1	4	0.4
Number of "Rights"	1186.	....	66.5
Number of Pupils	283.	....	283.
Mean	4.19	....	0.24
Median	4.7	....	0.
Tentative Standards 4.5-months, 2.8; 9-months, 5.6			

Discussion of Results

Validity.—The results for all the tests were carefully examined in order to determine their validity in measuring achievement in first-year algebra.

1. The problems in each of the three Hotz Scales are supposed to be arranged in order of difficulty. For the most part they were found to be so arranged. The problems in the Rugg-Clark tests were found to be less accurately evaluated in difficulty. Figures 1, 5, 6, and 7 represent graphically the variation in order of difficulty.

2. As to distribution of achievement of pupils tested, it was found that the three Hotz Scales gave fairly normal distributions of scores. The distributions in the Rugg-Clark problem and graph tests were irregular. The median score of zero made by the 283 pupils given the Rugg-Clark Graph Test, considered in connection with the Hotz results, justifies the conclusion that this test did not provide a valid measurement of graphing ability.

3. The validity of scores based upon the number of problems solved in the Rugg-Clark Formal Tests is not certain. The pupils were not uniform in parts of cycles solved. The problems are of unequal difficulty and are not arranged in order of increasing difficulty.

4. A detailed study of the ranking of both classes and individuals revealed the fact that there was less variation for scores based on the established value of each problem than for scores computed in terms of the number of problems solved.

5. A careful examination of the Rugg-Clark Tests 10, 11, 12, and 13 revealed many variations from the "cycle" principle as described by the authors. To illustrate: Test 12 consists of seven cycles of four problems each. Types 1 and 4 do appear in the same order in the several cycles. Type 3 is the same for cycles 1 and 7. It appears in the second place in cycles 2 and 4; is not found at all in cycles 3 and 5; and appears with one change of sign in cycle 6. Type 3 of cycles 2 and 3 is the same. It appears also in the second place in cycles 3, 5, and 6. Type 4 is the same in all the cycles; but the same type is found also in the third place in cycles 4 and 5 and in the second place in cycle 7. On pages 69 and 70 of their monograph,[1] the authors assert that

---

[1] H. O. Rugg and J. R. Clark: ***Scientific method in the reconstruction of ninth-grade mathematics.*** **(Supplementary Educational Monographs, University of Chicago Press, Vol. II, No. 1) 1918. 189 pp.**

failure to follow the principle exactly causes "quite different percentages of failure on the part of pupils solving the examples."

Economy.—A comparison of the tests with respect to economy of time showed:

1. The Rugg-Clark Formal Tests were the more burdensome to give and to take. This set of tests required ten minutes the first day and practically the entire period the second day.

2. The time required to score the papers received from the two problem tests was approximately the same. For the equation and formula tests approximately three times as much time was required to score the Rugg-Clark papers as to score the Hotz papers. No comparison could be made for the graph tests on account of the small number of problems solved in the Rugg-Clark graph test.

Difficulties.—Some of the difficulties encountered in the use of the tests were:

1. No directions are furnished by the authors for giving or for scoring the Rugg-Clark problem test.

2. Scoring in partial credits is more difficult and the fairness to pupils may be mainly apparent. It is not clear that the two parts of any of the problems are of equal value.

3. In the Rugg-Clark rate tests pupils are not instructed by the authors as to the form in which to leave answers. This lowered scores made by some pupils and caused difficulty in scoring.

4. The use of the word "equation" in problem 5 of the Rugg-Clark problem test caused difficulty. Some teachers were of the opinion that the use of the term "roots" in the Rugg-Clark Graph Test was the cause of failure on the part of some pupils to solve the problems in this test.

Diagnosis—

1. For diagnostic purposes no important difference between the problem tests was apparent. They represent about the same variety of algebraic statements and equations.

2. The Rugg-Clark Formal Tests furnish a greater number of problems of various types of equations than does the Hotz Equation and Formula Scale, and provide more diagnostic information regarding ability to solve algebraic equations.

3. The Hotz Graph Scale gave information as to ability to read graphs, to locate points when the coördinates are given, to find the coördinates of points, to plot linear equations and to solve problems by the use of graphs. This was done through the use of separate questions. The information furnished was thus of definite diagnostic value. The Rugg-Clark Graph Test is a test in finding roots by drawing graphs of simultaneous equations. The several steps involved in the process are not separated. Hence little explicit diagnostic evidence was obtained from this test.

4. Before any of the Hotz Scales can furnish much diagnostic evidence, pupils must have studied algebra at least a few weeks. Thus such evidence comes late to be of real use for needed changes in teaching emphasis. The Rugg-Clark Formal Tests may be given when the various types of equations are being studied. For such use it would be better if these tests were printed on separate sheets instead of in book form.

CLASSIFICATION OF KINDERGARTEN CHILDREN FOR FIRST GRADE BY MEANS OF THE BINET SCALE

CHARLES D. DAWSON
Assistant Superintendent of Schools, Grand Rapids

In the fall of 1920, at the request of superintendent W. A. Greeson of Grand Rapids, Michigan, Dr. Charles Scott Berry of the University of Michigan was secured to train approximately sixty kindergarten teachers and their assistants to give the Binet test. Through the united efforts of these sixty teachers, assisted by the kindergarten supervisor, Miss Annie J. Blanchard; her assistant, Miss Elizabeth Webster; and Mrs. Cordelia Creswell, supervisor of mental testing, 2029 kindergarten children were tested by June, 1921.

After the testing had been completed, the mental ages were reduced to a September-first basis and classified into three groups as follows: (1) in the X or highest group children seven years or older mentally; (2) in the Y or middle group children six to seven years old mentally; (3) in the Z or lowest group children five and a half to six years old mentally.

Table I shows the distribution of the 2029 kindergarten children according to mental and chronological ages.

TABLE I. THE DISTRIBUTION OF 2029 KINDERGARTEN CHILDREN ACCORDING TO MENTAL AND CHRONOLOGICAL AGES

Chronological Ages	Mental Ages					
	Below 5½	5½-6	6-7	7 and above	Totals	Per-cents
Below 5½..........	33	29	41	13	116	6
5½-6..............	79	100	255	111	545	27
6-7................	103	145	530	399	1177	58
7 and above.........	55	23	58	55	191	9
Totals.........	270	297	884	578	2029	
Percents.......	13	14	44	29		100
Groups........	Sub-Z	Z	Y	X		

It was recommended that the principal and the kindergarten teacher, after taking all factors into consideration, should either place the children who did not test five and a half years according to their mental age on September first, in one of the groups or leave them in the kindergarten. The kindergarten teacher was also privileged to reclassify a child whose mental test score placed him in a group to which, in her opinion, he did not properly belong.

THE WORK ASSIGNED TO EACH GROUP IN THE FIRST GRADE

All groups were to do the same type of work for the first ten weeks of school, i. e., regular first-grade work. This was done in order that first-grade teachers' judgments could be utilized in the experiment. During this probationary or trial period, the first-grade teacher was privileged to move a pupil from one group

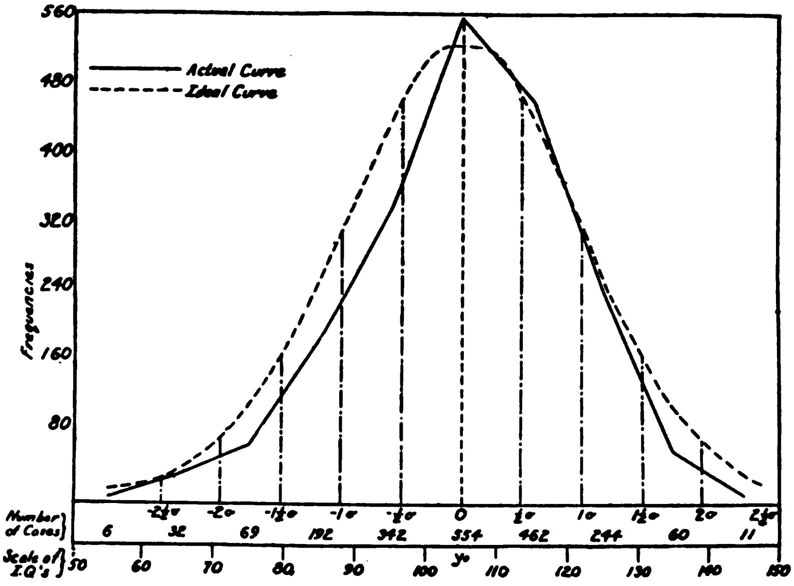


FIGURE 1. THE IDEAL CURVE SUPERIMPOSED ON THE ACTUAL CURVE OF DISTRIBUTION OF THE INTELLIGENCE QUOTIENTS OF 2029 KINDERGARTEN CHILDREN OF GRAND RAPIDS, MICHIGAN. TESTED JANUARY TO JUNE, 1921.

to another whenever in her judgment the child did not properly belong in the group to which he had been originally assigned. After ten weeks were over, or even before, if the first-grade teacher had settled definitely on the members of each group, it was recommended that the Z group go more slowly, doing *at least* a half-year of first-grade work in a year and as much more as the group was able to do. The other two groups were to do regular first-grade work, enriched as much as seemed expedient.

ACTUAL CLASSIFICATION

Even though 2029 kindergarten children were tested for first-grade work it did not prove feasible to form X, Y, and Z groups for each school. In schools in which there was only one first-grade teacher it was not deemed advisable to form three groups from their small number of pupils. Some of the teachers in the smaller buildings formed two groups only, an upper and a lower, designated A and B. The reason given for not forming three groups where there was only one teacher in the first grade, was that to do so would cut down the actual recitation periods of each group 50 percent. For example, three groups reciting their reading lessons for 20 minutes each consume 60 minutes. Were these same pupils divided into two sections, each section could continue for 30 minutes with no longer school day. That is, two 30-minute periods equal three 20-minute periods.

In schools having two first-grade teachers, the pupils could have been divided first into three groups, X, Y, and Z. One teacher could then have taken the X's and high Y's, while the other could have taken the low Y's and all the Z's. Unfortunately, or otherwise, this arrangement was not followed by all the larger schools. Some assigned three groups to each teacher, when it was possible for them to have formed two groups only for each teacher.

In consequence of the above optional types of classification, only 1129 or approximately 64 percent of the first-grade children were actually classified on the X, Y, and Z basis. Table II shows the actual classification of these children.

There were, however, 213 of these 1129 pupils who were not classified according to their mental ages but were placed in groups X, Y, and Z according to the judgments of the kindergarten teachers. The numbers representing these pupils are italicized in Table II. Of the 71 children who were below the mental age of the Z group, 64 were placed by the teachers in Z, 6 in Y, and 1 in X. Of those having the mental age of the Z group, 43 were placed in Y; while of those having the mental age of the Y group, 27 were placed in Z and 45 were placed in X. Twenty-seven of those having the mental age of the X group were placed by the teachers in the Y group. The remaining 916 or 81 percent were placed in the X, Y, and Z groups according to their mental ages.

STABILITY OF GROUPING

During the first ten weeks it was necessary to alter this original grouping to some extent. Of the 64 pupils having a mental age of less than five and a half years, who were placed in group Z, 24 were transferred to Y, or up a group, while 1 of the 6, placed originally in the Y group, was transferred to Z, or down a group. The changes within each group are recorded in Table II. Of the 213 pupils placed according to the kindergarten teachers' judgments, 25 were later transferred to a lower group, while 35 were placed in a higher group. That is, 153 or approximately 72 percent of them remained in the group in which the kindergarten teachers had placed them. In other words, approximately three out of every four remained unchanged.

Of the 916 placed in groups according to their mental ages, 56 were transferred to a higher group; while 82 were changed to a lower group, making a total of 141 transfers. Therefore, 775 or 85 percent of these remained unchanged. If we consider the number of changes within each group we find that 22 percent of the Z group were subsequently changed to another group; while 17 percent of those originally placed in the Y group and 16 percent of those in the original X group were likewise transferred.

TABLE II. THE CLASSIFICATION OF 1129 KINDERGARTEN CHILDREN ON THE X-Y-Z PLAN FOR FIRST-GRADE WORK

Mental Ages	Groups									Totals (Original Grouping)	Per-cent	Total Numbers Transferred
	Z			Y			X					
	Original number	Number transferred to groups:[a]		Original number	Number transferred to groups:[a]		Original number	Number transferred to groups:[a]				
		Y	X		X	Z		Y	Z			
Below 5½ years.....	*64*[b]	*24*	*0*	*6*	*0*	*1*	*1*	*0*	*0*	71	6	25
5½-6......	109	15	0	*43*	*3*	*9*	0	0	0	152	14	27
6-7........	*27*	*2*	*2*	499	41	38	*45*	*11*	*1*	571	50	95
7 and above	0	0	0	*27*	*4*	*3*	308	41	3	335	30	14
Totals	200	41	2	575	48	51	354	52	4	1129	100	198

[a] The changes indicated refer to shifts which were made during the first ten weeks of the first-grade work.

[b] The numbers in italics refer to pupils who were placed according to teachers' judgments.

GENERAL STATEMENTS CONCERNING KINDERGARTEN TEACHERS' JUDGMENTS AND TEST SCORES

Some are likely to conclude from these data that kindergarten pupils are grouped better by means of the Binet test scores than by kindergarten teachers' judgments because 85 percent of those placed in groups according to their test scores remained unchanged, while only 72 percent of those placed in groups other than where their test scores should have placed them, remained constant. It is unfair to draw such a conclusion from these data because in all cases the kindergarten teachers' judgments were utilized, for it was with their approval that a child was placed in the group which accorded with his mental age. Doubtful children only were placed in groups other than the one indicated by their test scores. If a kindergarten teacher felt that a child was not strong enough to keep up with the X group, even though he received a score which would place him in the X group, she recommended him for the Y or Z group. If she felt that a child had a language difficulty and did not test up to his ability she oftentimes placed him in a group above the one indicated by his test score. Other exceptional cases were similarly handled.

The conclusion which can be drawn from these data, however, is that it is remarkable that the Binet test, which takes approximately only 40 minutes to give to a kindergarten child, can so accurately group children according to their capabilities. Since only 198 out of 1129 cases were changed during the first ten weeks, there were approximately only 18 percent of the children who were readjusted. In other words, four out of every five were accurately placed according to their original classification.

REASONS FOR TESTING IN THE KINDERGARTEN

In the main, there are three reasons for testing kindergarten children for purposes of classification in the first grade:

1. The percent of failures in the first grade is very large. In fact, it is from two to three times that in any other elementary grade.

2. It is at least suspected that we are asking a great many children to do first-grade work who are not sufficiently developed mentally to do so.
3. First-grade teachers need assistance in diagnosing children's difficulties since the children come to them unsorted.

A report from first-grade teachers, relative to the original classification and changes made by them during the first ten weeks was requested. One important question asked, was, "Do you feel that this grouping by means of the Binet-test results has been of assistance to you? If so, in what way?" Forty-three of the fifty-eight first-grade teachers gave positive answers in favor of it; fifteen left the answer blank; while none gave a negative answer. Granting that all the fifteen who left the answer blank were not in favor of it (which is altogether improbable) the fact still remains that approximately 75 percent favored it.

Some of the typical answers made by first-grade teachers concerning the testing were:

"I feel it has been a great benefit by throwing light on what we should expect of children. If a child tests low, we know we must allow him more time to complete a grade. If he tests normal, but is not doing good work, it should be our duty to find out why and urge him to do better. If he tests superior, he should be given a chance to make the next grade with a little extra help."

"Yes, I think the grouping by means of the Binet test is most valuable. It gives the teacher some definite knowledge of the child and where to place him. This, of course, saves much time as well as giving the valuable information which helps so much in knowing each individual child."

"This means of grouping has been of assistance in the discovery of the very slow pupils before they become bewildered by trying to keep up with the brighter children. On the other hand, it does not subject the bright child to the dulling effect of waiting when a slow pupil is up for recitation."

"Yes, we have two 1-1 rooms at——————. For several years we have aimed at the same grouping that we have this year viz.: X and high Y pupils in Miss ——————'s room and the lower Y and Z pupils in my room. This year there has been less need for transfer from room to room. While I have found considerable variation within the assignments made to me as Y and Z groups, the X pupils seem to have been well sorted out."

"It has been of great assistance to me. Each year I have had sent to me for the advanced group children in whom the kindergarten teacher had mistaken manual dexterity for mental ability. Shyness is also often taken for slowness."

"Yes, because the children are better classified."

"I feel that this grouping has been of assistance to me. It has saved much time in having a definite organization of pupils at the beginning."

"At least a week of time was saved at the beginning of the term which is usually used in the adjusting of classes. Children who need special attention are brought to notice at the very beginning of the year."

"Yes, it is much easier to handle a large number in a class if they are of nearly the same type mentally. Because of this grouping we are able to work with two groups instead of three or four as formerly in first-grade rooms."

Advantages Summarized

1. Considering the report as a whole there seem to be several advantages to be gained from this testing. Great assistance is furnished the first-grade teacher in analyzing the child. Mere personal opinion is not always a safe guide in the analysis of children's abilities to do first-grade work for there are children who are brilliant in a general sense but dull in a technical sense. That is, by travel, picture-books, magazines, and fine home training they have received a general training far beyond that of children who are much older than they. But when it comes to teaching them that "c" should sometimes be given the sound of "k" and sometimes the sound of "s;" or that 9 and 5 are 14, etc., they fall far short of their teacher's expectations. On the other hand, many children are brilliant in a technical sense but dull in a general sense. That is, through lack of picture-books at home, magazines and travel, combined with indifferent home-training their limited experience with things in general makes them appear dull. But when asked to read page four in their reader, they can not only do it well but they can also reproduce every essential idea of the page, or if asked to give the sum of any of the forty-five combinations, they can do it without hesitation.

2. It is possible by means of the testing to sort out with a high degree of accuracy the children who are below five and a half years mentally and to permit only those who are healthy and seem mentally strong to go on to the first grade.

3. Much time is saved in the grouping of the children at the beginning of the semester. Almost all the experimental grouping is eliminated. Several teachers say that from a week to ten days is saved in the organization of their classes.

4. The teacher has an impersonal check on the teachability of the child. It sometimes happens that children who are teachable will not learn. Some children see no reason for learning to read. Their wants and pleasures are satisfied without knowing how to read, and unless the work can be motivated they put forth little or no effort to master the abstractions of reading. The following incident which happened in one of our schools illustrates the point. A short time ago a first-grade boy who had shown by his test that he was a perfectly normal boy was not learning to read. The teacher had used all her powers in attempting to induce him to try. The principal was called upon for assistance. When asked why he didn't want to learn to read, the boy replied that it would not be necessary for him to know how to read at the work he was going to do when he grew up. When asked what he was going to do, he said he was going to be a clerk in a grocery store like his father. But the principal said, "You will need to know how to read the orders brought in by little boys and girls to your store and also orders which are taken over the telephone, won't you?" The boy was told to talk the matter over with his father at home that evening. On returning to school the next morning he said, "Well, if that's the way of it, I guess I'll learn to read,"—and he did. The principal, and the first-grade teacher of this school, understand the meaning of Binet testing; and through test results they are able to distinguish between apparent and real difficulties. The impersonal feature of the test proves to be a "life-saver" in many cases.

5. A further advantage is to be found in having groups mentally homogeneous to begin with. It goes without saying that when first-grade pupils are grouped according to a teacher's judgment, the children are grouped mainly on performance and effort. When children of similar mentalities are grouped together, the

teacher feels that she should do all in her power to bring out a performance commensurate with the child's mentality before recommending him to a slower group. The child, therefore, has the proper chance with his mental equals in beginning his school life.

6. When there are two first-grade teachers in a building the plan of cutting down the number of groups for each teacher can be used. That is, it is not necessary for each teacher to have a fast, middle, and slow group. One teacher can take the fast and a part of the middle group while the other can take the other part of the middle group and the slow. In this way each teacher will have but two groups.

7. Possibly the greatest advantage lies in the fact that pupils in the first grade may be advanced as rapidly and only as rapidly as their mental development permits. This is not exactly true with each individual, but it is true with respect to the three groups as wholes. In other words, there is opportunity for enriching or accelerating the work of the X group, commensurate with its mental age; of doing ordinary work with the Y group; and of going slowly enough with the Z group so that the amount of work the pupils cover in a semester is within their grasp. Therefore, from the standpoint of habit formation in the thinking processes both as to speed and thoroughness, and from the standpoint of hygienic mental training and development from every angle, children are fed on an educational diet which agrees with them and enables them to grow naturally. In other words, the bright child is not held back and the dull one pushed ahead as is usually the case when children are grouped at random or by unscientific methods. If it be true, as some psychologists say, that bright children oftentimes have fixed their habit of reading as to speed, eye-movement, etc. by the time they have finished the third grade, then great injustice is being done to bright children all over the country since they are forming the habits of ordinary children because they are trained only as ordinary children. The ideal school would be one in which every child progressed according to his own mental level of intelligence and rate of mental

development. This, at present, is impracticable, especially if carried out in its minutiae, since such an organization would undoubtedly entail large expense. The Winnetka schools, near Chicago, are trying out this experiment.

A step can be taken in that direction, however, by grouping together children who vary no more than a year or a year and a half in mental age, instead of following the usual practice of putting children into the same class who differ by as much as three or four years in mental abilities.

LETTER-WRITING: A CURRICULUM STUDY IN ENGLISH COMPOSITION[1]

ROY IVAN JOHNSON
Stephens College, Columbia, Missouri

If we assume that the functional viewpoint in the construction of a curriculum is the correct one, then it becomes one of the objectives of the course in English composition to develop the ability to write a good letter, since the letter is the commonest form of written expression employed in the activities of a normal life.

But what is a good letter?

No satisfactory standard of measurement has been evolved. Some individual opinions have been hazarded and certain desirable characteristics have been set down in casual and isolated references. But, in general, letter-writing as an art has not been seriously studied with a view to teaching it in the schools.

This is partly explainable by the fact that a general skepticism prevails, or has prevailed, concerning the possibility of discovering a "technique of letter-writing" which can be adapted to instruction. "Personality is the secret of a good letter" has been the aphoristic declaration with which any proposal to teach letter-writing has been met. And the statement has been generally accepted as a comfortable solution of what otherwise would be a new English problem. But the art of letter-writing, like other arts, is a mode of self-expression and should be subject to effectual and helpful analysis. It is true that it is more intimate, so far as individual relationships are concerned, than painting or poetry; and, for that reason, the contention that it becomes the special vehicle of personality seems to have some logical basis. The argument, however, is not wholly tenable. There is often as much affectation among friends as before the public. The personality of Keats pervades *Endymion* as thoroughly as it does his letters. But even if one fully admits the personality argument, a careful study proves itself fruitful through the discovery of certain common characteristics and devices employed to advantage by expert users of the art.

---

[1] This study was conducted by the Stephens College Department of Research under the direction of Dr. W. W. Charters.

The present study was undertaken with a view to establishing well-defined, particularized objectives in the teaching of letter-writing to college students—that is, to discovering by the process of analysis and the compilation of expert opinion a "technique" of the art by which the general quality of letter-writing may be improved.

Successful teaching, however, depends not only upon a consciousness of an objective but also upon the proper determination of the subject matter. What shall be taught about letters?[2]

The materials of this study consist of: Group I, approximately one hundred and fifty letters contributed by women of recognized leadership throughout the country[3]; Group II, modern collections of literary letters; Group III, one thousand letters from young women of approximately college freshman standing.[4]

The analysis of the first group of letters revealed the prevalence of certain common characteristics. The analysis of literary letters contained in modern collections (Group II) lent validity to the findings of the first analysis by discovering substantially the same characteristics as those found in the letters of Group I—in addition to certain niceties of literary style which many of the letters of the first group did not exhibit. The analysis of the third group discovered certain general weaknesses of student letters and established the "mechanical" needs of college students so far as letter-composition is concerned. It also made possible the determination of a norm of accomplishment which could be used as a basis of comparison in judging other groups of student letters.

---

[2] An examination of fifty textbooks in composition and rhetoric revealed the fact that less than four percent of the space in texts is devoted to letter-writing—and this four percent is almost entirely given over to the *forms* of business correspondence and to formal notes. A diligent search through the pages of these texts disclosed only casual references to *content* characteristics of the letter. Such adjectives as *clear, tactful, courteous, sincere, unstudied, vivid, full of life,* and *pleasing* are used to describe the style and quality of good letters. Save for these brief and incidental records of an author's individual opinion, little help for the amateur letter-writer is to be found in the texts except in matters of form and grammatical correctness.

[3] These were current letters from the private correspondence files of well-known women. The letters were contributed as samples of "good" letters. They were very kindly offered in response to requests for cooperation and were of invaluable assistance in the study.

[4] Two hundred and fifty of these letters were social letters and were obtained through the confidential cooperation of one hundred and fifty of the freshmen English students at Stephens College. Seven hundred and fifty were business letters obtained from the college office.

CHARACTERISTICS OF GOOD LETTERS AS DETERMINED BY AN ANALYSIS OF GROUPS ONE AND TWO

It is not to be supposed that the characteristics discussed in the following sections were found in every letter which was examined, but they recurred with noticeable frequency throughout both groups. A count was taken on a random sampling of the letters and the frequency of each characteristic was found to be at least one-half. In other words, each characteristic here discussed occurred in fifty percent or more of the letters examined.

Courtesy.[5]—In intimate or friendly letters this amounts to much more than formal politeness. It implies thoughtfulness, solicitude, consideration for others, a well-bred deference to the epistolary rights and interests of the person to whom the letter is directed. Some of the devices, or ways, by which that attitude of mind which is here denominated as courtesy may be expressed in a letter are[6]:

1. Give due attention to the points in the letter which is being answered.
2. Never fail to express expected congratulations.
3. Never fail to acknowledge obligations. Show a sense of gratitude.
4. Express good wishes.
5. Make inquiry about matters which are of the first importance and interest to the recipient of the letter.
6. Apologize, if there is any occasion for it.
7. Express or imply respect, esteem, or affection.
8. Adapt subject matter to the age and interests of the reader (e. g., letters to children).

---

[5] These characteristics apply primarily to social correspondence since the study is made largely in terms of the friendly, or social, letter.

[6] All devices listed were utilized in letters examined—not all of them, of course, in any one letter.

Informality.—One naturally expects informality in social letters. But it requires a nice sense of discretion to be informal without being "cheap." Informality breaks down distance and insinuates a comfortable sense of nearness. The indulgence of a tendency to be natural plus a certain caution against inelegant expressions should produce a desirable informality.

"An odd leaf of a volume, and that volume your friend," says Thoreau in characterizing one of his own letters. But that odd leaf must bear the scrawls and pencilings, the sly smiles and whimsicalities, dictated and inspired by a human soul off-duty in an attitude of unaffected comfort. Naturally there are times when the conventional asserts itself. In fact it may dominate the style of the letter without destroying certain touches of informality that characterize occasional sentences. Perhaps this element of informality had better be called a "flavor" rather than a style of composition characteristic of an entire letter or series of letters.

In the letters of Group I certain interesting "ways" of being informal were observed:

1. Use of colloquialisms, or conversational expressions.
2. Use of invented expressions.
3. Use of pen sketches and illustrations.
4. Elimination of over-formal headings and conclusions.
5. Use of figures of speech humorously exaggerated.

Humor.—This characteristic is possibly allied to informality. But it is certainly evident both in the published letters and in those collected from the personal correspondence of prominent women. The familiar letter affords an excellent channel for humorous expression. In a mood of delightful relaxation, the letter-writer often becomes a mental contortionist, producing smiles by rapid revolutions of thought, by exaggeration, or by some impossible comparison. Following are some of the common sources from which letter-writers draw their humor:

1. They smile at the expense of human weakness (sympathetically, of course).
2. They depict absurd situations.
3. They exaggerate—intentionally.
4. They juxtapose apparently unrelated thoughts.

5. They make reference to amusing experiences.
6. They indulge in plays upon words.

Humor, as a condiment of letter-writing, is evidently important if one may judge by its prevalence in the letters studied. Its use is not restricted to the letters of the professional humorist—whose entire letter may billow with laughs—but it is freely and liberally used (even in letters of serious purpose) by Lamb, Stevenson, Moody, and practically all writers of pleasing letters.

Opinions.—The letters of Groups I and II were distinctly opinionative. In one of Bayard Taylor's letters he apologizes because the letter contains so little besides news. And Christina Rosetti says: "News is not my forte." Even in those letters (like George Meredith's) which seem to be especially "newsy," the thing that gives life to the letters is the running comment—the exposition or implication of attitude. A suggestive and illustrative passage from Meredith's letters may be cited here. In telling of a tramp across the hills with some of his eminent friends, he writes: "They are men of distinction in science and literature. Tramping with them, one has the world under review as well as pretty scenery." While it is not necessary for one to have "the world under review" in the letters he writes, it *is* desirable that he show an intelligent responsiveness to that small part of it which has come under his direct observation. Opinionative letters are intellectual reactions toward life. A letter which is merely "newsy" may be only a bulletin of events.

The best letters are likely not to be descriptive[7] or wholly narrative, but they will challenge the interest by an exposition of attitude or a casual interpolation of opinion relative to matters that concern (either directly or indirectly) one or both parties to the correspondence. It is well, however, to keep in mind Mr. Clarke's caution (Clarke, Thomas Arkle. *When You Write a Letter*); viz., that a letter should not be a philosophical essay.

---

[7] Very little description was found in the letters studied. This may be due to the fact that description is a very difficult form of discourse for most people to use clearly and effectively. Or it may be due to a feeling on the part of writers that too much description "clogs a letter" (as Wordsworth says).

For it may be that in attempting to write "opinionative" letters, we shall turn serious-minded and approach our task like the plenipotentiaries of world powers negotiating peace or war.

Let it be understood that the "newsy" letter is not taboo. To convey news—well-sorted news—is one of the motives of correspondence. But ordinarily news should be considered simply as subject matter in the rough. It may be refined by proper organization, by opinionative comment, by humor, etc. But without embellishment it becomes the dry rot of social letter-writing.

Optimism.—It has been remarked that the letter is a vehicle of "fugitive moods."[8] It is a verbal chameleon against the changing hues of human temperament. The general prevalence of optimism in the letters of Groups I and II of this study may be ascribed to the fact that optimism is naturally the dominant mood of the normal mind. A few of the letters, however, are distinctly dispirited. While there is, apparently, a place for the letter expressive of various moods—moods of cheerfulness or discouragement—it may be set down as a safe conclusion that a courteous consideration of the feelings of the reader demands only a rare indulgence in despondency. The optimistic letter has immeasurably greater possibilities of winning interest and giving pleasure than the letter of deep-tinted "blues."

The appropriateness of the mood may depend upon the truth of the mood. True reactions of the writer often make expressions of discouragement necessary. There is an intimacy about the social letter which should forbid an *affectation* of optimism. If correspondence is to be promoted, real and truthful expressions of attitude are necessary. But assuming that cheerfulness is the *true* mood of normal life, it should find a large place in the friendly letter. At all events, the significant fact remains that a cheerful outlook on experience characterized practically all the letters of Group I—letters which were selected by competent critics as marking high standards of correspondence.

---

[8] Dawson, W. J. and C. W. *The great English letter writers.* Vol. 1, p. 31.

Few centers of organization.—Another conclusion reached through an examination of letters of Groups I and II is that good letters show remarkably few instances of loose organization. The amateur's letter may be a miscellaneous collection of paragraphs—a faithful record of disconnected facts. But the best practicers of the art favor the "linear" letter—the letter which confines itself to few and closely unified subjects of thought. It is conceivable that a "gossipy" letter might sometimes be of interest to a correspondent, but in such instances the choice bits become merely the illustrative and coherent development of the general theme of "gossip." Matters entirely foreign to the main theme of a letter may be left for a later communication, or thrown into a postscript, or treated in a very subordinate or incidental manner. The point is that the dominant theme of the letter should not be swamped with a flood of incoherent details.

Form.—Apparently there are almost as many variations of good form in letter-writing as there are good letter-writers. The most common form employed is herewith illustrated:

Columbia, Missouri,
September 22, 1922.

Dear Anna:
When I heard that your had won the prize offered
by ....
....
....

Cordially yours,
Mary K.

The heading, complimentary close, and salutation were found to change with the degree of intimacy or familiarity. The following salutations are arranged in what appeared to be the ascending order of formality: *My dear Anna, Dear Anna, Dear Miss Malvern, My dear Miss Malvern.* In very intimate letters, strikingly informal salutations are often found, such as nicknames, initials, etc. The complimentary close may also become less conventional, such expressions as *With love, Ever yours, Affectionately,* etc., being employed.

The more formal salutations in the letters were usually preceded by the name and address of the person to whom the letter was written; e. g.,

Columbia, Missouri,
September 22, 1922.

Miss Anna Malvern,
Paducah, Kentucky.
My dear Miss Malvern:
When I heard that you had won the prize offered by ....

In the replies of the women through whose assistance this study was made possible, twenty variations in form were noted in heading, salutation, and address. The most prevalent usage, however, conformed to the foregoing illustration.

One woman very well known in the field of modern literature writes that many "scrappy notes" and "illegible scrawls" which she has received from certain contemporary celebrities and which entirely ignore the matter of form are among the best letters she has received. She defines a good letter as "any letter which promotes correspondence between people." This is a thoroughly tenable attitude. Content and personal interest are greater factors in letter-writing than is the matter of form; yet most rhetorics insist upon inverting the emphasis. They teach arbitrary forms and ignore content.

However, the laws of form (including punctuation, capitalization, spelling, legibility, neatness, paragraphing, proper headings, etc.) are important and they must be taught. For the world at large is not so generous in its tolerance of "scrappy notes" and "illegible scrawls" as the woman, just referred to, who treasures them among her correspondence keepsakes. A fair conformity to such laws is to a letter what conventional manners are to a gentleman—a non-essential of character but an index of good breeding.

ANALYSIS OF STUDENT LETTERS

In regard to formal error.—A careful examination of two hundred and fifty social letters written by students (young women of approximately college freshman standing) was undertaken with a view to discovering the prevailing weaknesses in the letter-writing of those students who were ready for the college course in English. In this study two things were observed: (1) the amount of technical error found in the letters, and (2) the failure of the writers to employ interest devices or methods which were found to be generally characteristic of the "good" letters used in the first part of the study.

The categories of error used in the tabulations (see Table I) are practically the same as those established by a previous study.[9] The ease with which the errors lend themselves to this classification is the chief argument for the adoption of these categories for use in such studies as the present one. It also facilitates the comparison of results when such comparison is desirable.

In addition to the 250 social letters, 750 business letters were examined. These letters were written by high-school seniors who were considering entering college. The analysis of these letters furnished additional data concerning the distribution of

TABLE I. THE RELATIVE PREVALENCE OF ERROR IN 250 SOCIAL LETTERS WITH A TOTAL OF 90,530 WORDS

(106 *Unclassified Errors Not Included in this Table.*)

Error-Types	Number of Errors	Rank
I. Punctuation	889	1
II. Sentence Structure	517	2
III. Spelling	308	3
IV. Adjectives and Adverbs	269	4
V. Letter Form	223	5
VI. Capitalization	131	6
VII. Diction	97	7
VIII. Apostrophe	88	8
IX. Careless Writing	61	9
X. Verbs	59	10
XI. Prepositions and Conjunctions	47	11
XII. Case	25	12
XIII. Pronouns (not case)	22	13
XIV. Quotation Marks	9	14
Total	2745	

[9] See Johnson, Roy Ivan. "The persistency of errors in English composition," *The School Review*, 25:555 ff, October, 1917.

errors in student letters. Table II shows the error-types arranged in decreasing order of prevalence. (The table does not include 62 unclassified errors.)

TABLE II. THE RELATIVE PREVALENCE OF ERROR IN 750 BUSINESS LETTERS WITH A TOTAL OF 60,610 WORDS

Error-Types	Number of Errors	Rank
Punctuation	2242	1
Letter Form	402	2
Capitalization	273	3
Sentence Structure	263	4
Spelling	242	5
Verbs	222	6
Apostrophe	80	7
Diction	56	8
Careless Writing	53	9
Prepositions and Conjunctions	47	10
Adjectives and Adverbs	28	11
Quotation Marks	10	12
Pronouns (not case)	9	13
Case	7	14
Total	3934	

Examination of Tables I and II will show the following facts: (1) errors in punctuation constitute 45 percent of the total error; (2) the frequency of error in the business letters is approximately twice as great as in the social letters; (3) the category of error showing the greatest difference in rank in the two sets of material is "adjectives and adverbs."

The unusually large amount of punctuation error is partly explained by the determinative nature of the subject matter. Letter-headings, containing abbreviations, offer tempting opportunities for careless punctuation.

A possible explanation of the greater frequency of error in the business letters lies in the fact that there was a higher degree of selection in the social letters. Nearly all the business letters were written by girls who had *not* entered college—many of them never continued their training further than the completion of the high school—whereas the great majority of the social letters were written by students who had already entered upon their college course. Granting that the transition from school to college is automatically a selective process, *some* decrease in the frequency of error would be expected. The excessive amount of error in the business letters may also indicate a tendency toward carelessness in the composition of that type of letter. This conclusion is strenghtened by the fact that handwriting was much poorer in the business letters than in the social letters and that little attention was paid to the general neatness of the letter or to the kinds of stationery or writing materials used.

The colloquial nature of the social letter probably accounts for the drop in rank from *fourth* in Table I to *eleventh* in Table II of erroneous usage of adjectives and adverbs. "I sure want you to write" and "We had a real nice time" are expressions more likely to be called out in social correspondence than in formal business letters where there is less occasion for adverbial and adjectival emphasis.

In order to test the reliability of the data and determine the degree of general diffusion of error throughout the letters (i. e., regular recurrence of error-types) the social letters were divided into groups of approximately 18,000 words each and the frequency of error in each group was tabulated. In order to show the degree of constancy with which the error-types persist throughout the five samples, a table of ranks was prepared.

TABLE III. RANK OF ERROR-TYPES IN FIVE SAMPLES OF 18,000 WORDS EACH

Error-Types (from Table I)	Rank in 90,000 words	Rank in five divisional samples				
		1	2	3	4	5
I	1	1	1	1	1	1
II	2	2	2	2	2	2
III	3	3	3	4	4	4
IV	4	8	5	5	3	3
V	5	4	4	3	6	5
VI	6	5	6	9	5	9
VII	7	6	7	6	12	7
VIII	8	7	10	7	7	6
IX	9	9	14	8	9	10
X	10	11	9	10	8	8
XI	11	10	8	11	10	12
XII	12	14	11	12	11	11
XIII	13	12	12	13	13	13
XIV	14	13	13	14	14	14

By applying the formula for rank correlation to the data contained in Table III, correlation coefficients were obtained as shown in Table IV.

TABLE IV. CORRELATION OF RANKS OF ERROR-TYPES IN 90,000 WORDS WITH RANKS IN FIVE DIVISIONAL SAMPLES OF 18,000 WORDS EACH

Samples	Correlation
1	0.94
2	0.90
3	0.96
4	0.92
5	0.95

From Tables III and IV it appears to be a justifiable conclusion that the different types of composition error distribute themselves in approximately the same manner in small bodies of written material (about 50-pupil samples) as in large bodies of material (250-pupil samples). This proves the reliability of the present data so far as mechanical error is concerned and indicates the possibility of reducing the tabulation labor involved in other composition studies.

In regard to content.—When the social letters were examined from the standpoint of content and general style instead of from the standpoint of mechanical correctness, five obvious weaknesses appeared. In each case the characteristics of excellence established by a study of the letters submitted by prominent women (Group I) and of literary letters (Group II) have been used as a standard of judgment. It was not possible, however, to draw reliable conclusions in regard to courtesy in student letters. While there were few noticeable instances of *discourtesy* recognizable, the majority of the letters *seemed* to lack that thoughtfulness and consideration for the reader which was so obvious in the letters of Group I. A too frequent focusing on self (quite natural and excusable in youth) seemed to exclude too completely a consciousness of the interests of the "other person" and leave them at best only a peripheral margin of attention.

Of optimism there was no lack. The objection, if any, would be the seeming perversion of a healthful cheerfulness into an unballasted hilarity—a superlative gush—innocent of all adequate consideration of fact and circumstance.

Crudeness in letter-writing is the miscarriage of informality. The use of a slangy phrase that offends the sense of propriety, the continuous use of beheaded sentences, the excessive use of nicknames and familiar epithets, the hilarious and unrestrained treatment of personal experience, and over-indulgence in the "emotional superlative," may spring from a misdirected attempt to achieve informality in letter-writing—and there can be no doubt that informality is the result—but it is an informality not pleasing to a discriminating mind. This is the first notable weakness of student letters. Evidently the cure lies in the development of good taste, a sense of fitness, and in the careful study and application of the common devices of informality used by good letter-writers.

Inferior humor was evident in most of the student letters in which humor was attempted. The two most prevalent humor devices employed were: (1) interpolated laughter, and (2) the comedy of "tragic" occurrences such as the failure of recipes or of attempts to foil the teacher. Many of these instances really contained humor possibilities which were ruined in the telling. Interpolated laughter is, of course, inane and intolerable. A few rare instances of it can be found in literary letters which are otherwise unobjectionable. The improvement of the quality of humor employed in student composition can be accomplished through more definite attention to the subject of humor. A sense of humor needs educational direction as much as the capacity for musical appreciation.

Few opinions, other than exclamatory outbursts indicating approval or disapproval, were found in the student letters. Actual count showed fewer than ten percent of them to be opinionative. They were primarily "newsy," a complete gossip repertory, a catalogue of "she-did's" and "we-went's."

Lack of organization about central themes was another prominent weakness of the student letters. Instead of few centers of organization, there were many. The tendency to "tell everything" is probably a natural juvenile trait, since everything in one's experience is invested with a certain interest; but it is contrary to the best practice as found in the letters of Groups I and II. The value of *dominant themes* in letters should be stressed.

Violations of good form include poor paragraphing and improper headings and conclusions. Approximately one half the errors in form which were found in the student letters were errors in paragraphing. In listing improper headings and conclusions a wide latitude of usage was allowed. Strictly speaking, the matter of form should not be included in this list of weaknesses, since it is *not* a matter of content. But since paragraphing is so intimately related to the manipulation of the subject matter and is determined directly by the nature of the content, it is included here even at the risk of seeming illogical.

Summary

There may be other characteristics of a good letter than those which have been pointed out in this study, but the importance of the study lies in the fact that it establishes (by careful analysis rather than by mere opinion or guess) certain very clear objectives to be used in a course in letter-writing. It also records certain devices or means by which successful letter-writers have attained the objectives of courtesy, informality, humor, etc.

The analysis of student letters lends new confirmation to the belief that emphasis upon mechanical correctness is a primary need in student composition. It also shows the greatest weaknesses of student letters (so far as actual content is concerned) to lie in crude attempts at informality and humor and an excessive newsiness which expresses itself in "an abundance of detail without organization."

This study in no way minimizes the element of personality in letter-writing. In accepting the objectives which the study establishes, one is not "reducing personality to mathematics."

There must be a technique of every art, and letter-writing is no exception.

Having particularized the objectives of letter-writing, one naturally seeks the means of attainment. There arises, then, the need of a body of subject matter, based on the results of the present study or of similar studies and organized for use in the school course in composition. Not only is such a procedure possible in letter-writing, but it is possible also to study every other unit of functional composition and to determine scientifically the nature of the subject matter necessary to the development of proficiency in the specific task to be mastered.

DETERMINING CHRONOLOGICAL AGE IN DECIMAL PARTS OF A YEAR

HERBERT A. TOOPS
Teachers College, Columbia University

In the determination of age norms of tests, it is very desirable to take accurate account of the subjects' ages. Chapman[1] has recently called attention to this important matter. Inaccuracies in age determination may be largely avoided by asking the following three questions on the questionnaire filled out by each child:

1. In what year, month, and day *were you born?*

....
(Year) *(Month)* *(Day of Month)*

2. How old were you at your LAST birthday?..................................
(Years old)

3. What was the date of your LAST birthday?

....
(Year) *(Month)* *(Day of Month)*

It will be seen that the answer to 3 minus the answer to 2 should be exactly the answer to 1. Where all three of these do not check, personal inquiry of parents may elicit the required information.

Having presumably accurately determined the answer to question 1, it is still necessary to determine the exact age of the subject *upon the given test date* in order to place him in the proper fractional year-group for the determination of norms. This is readily accomplished by means of the accompanying table, made up on the assumption of 365 days in the year, so that 1 day = 0.0027397 years. By subtraction of any two entries of this table, the true interval of time elapsing between the two dates of the entries is found, correct to the one-thousandth part of the year. The table is made out for 12:01 A.M. of the dates indicated; so that for the additional fractional parts of a day one must add the corresponding fractional part of 0.0027397 to the interval found for corresponding clock times of the two days under consideration. The use of the table will be amply illustrated by the two examples:[2]

---

[1] Chapman, J. C. "Some elementary statistical considerations in educational measurements." *Journal of Educational Research,* 4: 212-220, October, 1921.

Example 1. Required the fractional part of a year between the birth date December 10, 1909, and the test date May 3, 1920. Under May 3 we find the entry, 0.334; under December 10 we find 0.940. We are now ready to subtract, thus obtaining the difference of 10.394 years.

1920.334
1909.940

Difference = 10.394 years

A calculating machine is very efficient in obtaining chronological ages. The 1920.334 is a constant and may be left in the machine all the time, merely altering the keyboard to suit the subtrahend.

Example 2. In vocational tests, it is often desirable to get comparative measures of workers' production at a stated interval of time after entrance upon the job. This general problem may be accomplished by the table, although a calendar giving the ordinal numbers of all days in the year may be handier where small intervals of time are concerned.

Required to know on what dates A, B, C, respectively, will have been on the job just 10 weeks.

10 weeks = 70 days. In the table 1 day = 0.0027397; whence, 70 days = 0.192 yr.

	a	b	c	d
	Date Entered Job	Year and Decimal	Col.b + 0.192	Dates on Which to Begin Count of Production
A	1-19-1919	1919.049	1919.241	March 30, 1919
B	3-15-1920	1920.200	1920.392	May 24, 1920
C	10-24-1918	1918.811	1919.003	January 2, 1919

We simply add 0.192 year to the entering date, expressed as a decimal (Col. b), and look up the resulting sum (Col. c) in the table, giving the dates of Col. d, which are the dates of the beginning of the eleventh week on the job.

NOTE: In Example 1, if the extra day of a leap year in February is included, allowance may be made by *adding* 0.003 year. In Example 2, in the same situation, one day should be *subtracted* from the date of Col. d.

---

[2] Since writing this article it has come to the attention of the author that the Monroe Calculating Machine Company has published a table for determining the interval between interest dates, based upon the time at 12:00 P.M. In order to avoid possible copyright infringement, the present chart is published with the permission of the Monroe Calculating Machine Company.

TABLE I. FOR DETERMINING TIME INTERVALS IN DECIMAL PARTS OF A YEAR

(1 day = 0.0027397 years)

	1	2	3	4	5	6	7	8	9	10	11	12	
	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	
1	.000	.085	.162	.247	.329	.414	.496	.581	.666	.748	.833	.915	1
2	.003	.088	.164	.249	.332	.416	.499	.584	.668	.751	.836	.918	2
3	.005	.090	.167	.252	.334	.419	.501	.586	.671	.753	.838	.921	3
4	.008	.093	.170	.255	.337	.422	.504	.589	.674	.756	.841	.923	4
5	.011	.096	.173	.258	.340	.425	.507	.592	.677	.759	.844	.926	5
6	.014	.099	.175	.260	.342	.427	.510	.595	.679	.762	.847	.929	6
7	.016	.101	.178	.263	.345	.430	.512	.597	.682	.764	.849	.931	7
8	.019	.104	.181	.266	.348	.433	515	.600	.685	.767	.852	.934	8
9	.022	.107	.184	.268	.351	.436	.518	.603	.688	.770	.855	.937	9
10	.025	.110	.186	.271	.353	.438	.521	.605	.690	.773	.858	.940	10
11	.027	.112	.189	.274	.356	.441	.523	.608	.693	.775	.860	.942	11
12	.030	.115	.192	.277	.359	.444	.526	.611	.696	.778	.863	.945	12
13	.033	.118	.195	.279	.362	.447	.529	.614	.699	.781	.866	.948	13
14	.036	.121	.197	.282	.364	.449	.532	.616	.701	.784	.868	.951	14
15	.038	.123	.200	.285	.367	.452	.534	.619	.704	.786	.871	.953	15
16	.041	.126	.203	.288	.370	.455	.537	.622	.707	.789	.874	.956	16
17	.044	.129	.205	.290	.373	.458	.540	.625	.710	.792	.877	.959	17
18	.047	.132	.208	.293	.375	.460	.542	.627	.712	.795	.879	.962	18
19	.049	.134	.211	.296	.378	.463	.545	.630	.715	.797	.882	.964	19
20	.052	.137	.214	.299	.381	.466	.548	.633	.718	.800	.885	.967	20
21	.055	.140	.216	.301	.384	.468	.551	.636	.721	.803	.888	.970	21
22	.058	.142	.219	.304	.386	.471	.553	.638	.723	.805	.890	.973	22
23	.060	.145	.222	.307	.389	.474	.556	.641	.726	.808	.893	.975	23
24	.063	.148	.225	.310	.392	.477	.559	.644	.729	.811	.896	.978	24
25	.066	.151	.227	.312	.395	.479	.562	.647	.731	.814	.899	.981	25
26	.068	.153	.230	.315	.397	.482	.564	.649	.734	.816	.901	.984	26
27	.071	.156	.233	.318	.400	.485	.567	.652	.737	.819	.904	.986	27
28	.074	.159	.236	.321	.403	.488	.570	.655	.740	.822	.907	.989	28
29	.077	....	.238	.323	.405	.490	.573	.658	.742	.825	.910	.992	29
30	.079	....	.241	.326	.408	.493	.575	.660	.745	.827	.912	.995	30
31	.082	....	.244	....	.411	....	.578	.663	....	.830	....	.997	31
	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	

Editorials

THE MENTAL AGE OF ADULTS

When an important weekly journal devotes one of its main articles to the subject of the mental age of adults[1] we may fairly say that it is no longer a merely academic question. We may be pleased that our problems appear to non-technical observers to be important, but we are not flattered when a non-technically trained writer is assigned the task of instructing the public in a matter which one might suppose would require somewhat special knowledge for its adequate treatment. It must be admitted in the present case, however, that the non-technical student has put his finger on some aspects of the question which have been overlooked by our colleagues in psychology. While he supports his conclusion in a very vague and incomplete way, his main contention is correct. This may be shown by reviewing the arguments in the case.

The whole controversy started with the official report of the army mental examinations, in which we read: "It appears that the intelligence of the principal sample of the white draft, when transmuted from Alpha and Beta Examinations into terms of mental age, is about 13 years (13.08)."[2] The conflict between this estimate and Terman's former estimate of 16 years, which had been pretty generally accepted, is discussed as follows: "For norms of adult intelligence the results of the Army examinations are undoubtedly the most representative. It is customary to say that the mental age of the average adult is about 16 years. This figure is based, however, upon examinations of only 62 persons,

---

[1] Lippmann, Walter, "The mental age of adults." *New Republic*, 22:213, October 25, 1922.

[2] Yerkes, Robert M. (editor) *Psychological examining in the United States Army.* (Memoirs of National Academy of Science, Vol. 15) Washington, D. C.: Government Printing Office, 1921. p. 785.

32 of them high-school pupils from 16 to 20 years of age, and 30 of them 'business men of moderate success and of very limited educational advantages.' This group is too small to give very reliable results and is furthermore probably not typical."

Curiously enough Mr. Lippmann draws exactly the opposite conclusion from this discrepancy in results and difference in number of cases from that which is drawn by the writer of the army report. The latter holds that, because of the preponderance of evidence, the army figures may be taken as clear evidence that the average adult intelligence is a little over 13 years. Mr. Lippmann, on the other hand, points to the general fact that "the average of the army did not at all agree with the average of Mr. Terman's Californians" and that since this is so, and the army scores are based on so many more cases, the army tests have "knocked the Stanford-Binet measure of intelligence into a cocked hat."

Just how the army tests have discredited the Stanford-Binet norms we are not told except in vague phrases. The issue would seem to be this. Those who accept the adult norm of 13+ years regard the Stanford-Binet scale of norms as valid and then express the army median score in terms of these norms by a translation of Alpha and Beta scores into them. Mr. Lippmann's point would seem to be that this procedure contains a flat contradiction. If you accept the finding that the average adult has a mental age of 13+ years you cannot at the same time hold that the mental age of the average adult is 16 years — which is involved in the acceptance of the Stanford-Binet scale of norms. The Stanford-Binet scale must therefore be abandoned, since it was derived from so much the smaller number of cases. To use Mr. Lippmann's picturesque phrase, "it is knocked into a cocked hat."

There seems indeed to be much merit in this contention. How can we at the same time use the scale for the interpretation of the mental age of adults and deny its validity by adopting a mental-age standard which is at variance with it by three years? We cannot. That is plain.

But pause a moment. Are we then to consider the whole Stanford-Binet scale of norms as seriously in error? Surely that would not be justified by our widespread experience with it. Its essential correctness is certified not merely by the thousand or so children upon whom it was standardized but by the thousands who have been tested by it since it was standardized. If the whole scale is not in error, just where does it begin to be in error, at the age of 15, or 14? If it is correct up to 14 and in error by three years at 16 how does it happen that the tremendous gap in difficulty does not throw out the scores of children of 13 and 14?

There is a much clearer solution which avoids this dilemma. The median mental age of recruits was estimated by giving six hundred odd men both the Army Alpha test and the Stanford-Binet scale. From a comparison of scores a table of equivalent scores in the two tests was worked out. By means of this table the median mental age of recruits was calculated. *This procedure assumes that a test standardized with children will give equivalent scores with adults.* It assumes that the average man, who has been out of school for from five to ten years, who has done very little reading and has been provided very meager intellectual stimulus, will do as well on a test standardized on school children as he would have done while he was in school. There is good reason to believe that he would not do as well. The character and amount of schooling influence the score in mental tests. Lapse of schooling, therefore, must influence it.

The dilemma into which we are forced by this customary assumption that scores of adults can be translated into mental-age norms by means of tests standardized on children is obvious when we attempt to estimate the proportion of mentally defective persons. Terman has calculated that 2.3 + percent of children have an I.Q. of 75. This means that their ultimate mental age is 12 or below. Using the mental-age equivalents of the army scores, however, we arrive at 30 + percent as the percent of adults whose mental age is 12 or below. Are we to suppose that 27.7 percent have *become defective* between school age and manhood? Surely not. Yet there is no escape from this absurd conclusion if we grant the validity of the Stanford-Binet scale and of the translation method above mentioned.

The simplest way to put the question "What is the mental age of adults?" is "When do individuals cease to grow intellectually?" The best evidence that I have seen on this point is that mental growth, as measured by a mental-test score, continues well beyond 13 years of age, and probably beyond 16 years. This evidence consists of repeated measurements on the same individuals. Such measurements constitute the only thoroughly reliable evidence on the question.

Let us be clear on the main fact. Intellectual growth does not cease at 13 years and the adult does not possess the inherent intellectual limitations of the thirteen-year-old child. The attempt to make tests show that adult mentality is 13 gets us into contradictions. The extravagant applications which have been made of this supposed discovery are enough in themselves to make us sceptical of it. The error has become so widely diffused that it has become a matter of importance for us to take some pains to correct it.

F. N. F.

THE DIFFICULTY OF TEXTBOOKS

It is clear that the ability of children to read their textbook definitely conditions the value of the book. Many other characteristics must be considered in the selection of a textbook, but none of them is effective unless the book can be understood.

The appearance of Franzen and Knight's "Textbook Selection"[1] and their use of the difficulty criterion prompt us to comment appreciatively upon the value of their data. Our studies of the product of teaching have proceeded far enough so that we can state the reading ability of children in each grade. Moreover, we can do so in units which have the same meaning when applied to reading material. By bringing together the two ideas —namely the ability of the children of a given grade and the difficulty of the textbook placed in their hands—Messrs. Franzen and Knight are able to make some interesting statements concerning the difficulty of geography texts. For example, it is found that five-sixths of the material in the typical first geography book—the book usually begun in grade IV—is too hard for the pupils of that grade to understand. It is also stated that only about one-half of the second book of the usual two-book series, can be understood by the children of grade VI—the grade in which it is first introduced.

---

[1] Baltimore: Warwick and York, 1922. 94 pp.

Observant teachers have for a long time noted the discrepancy between the reading ability required by textbooks and the reading ability actually possessed by pupils. Vocabulary, style, and organization of textbooks have been such that they have not reached their proper reading public, namely, the children. These writers by experiment have confirmed this belief regarding geographies. There is every reason to believe that similar findings would be obtained with respect to textbooks in history, literature, general science, and arithmetic. The method of the investigation is simple and direct. It proceeds upon the sufficient basis of the selection of passages from the book under consideration and the formulation of questions on each passage. Certain precautions are of course to be used in making up the passages and questions. The real point, however, is that our statistical analyses—those despised meanderings of the theorists—have placed us in a position where we can work simply, directly, and conclusively to a determination which everybody can understand. A geography five-sixths of whose material cannot be understood is a tangible, as well as a disconcerting, piece of pedagogical luggage.

Either or both of two courses are open to us. The geographies—and inferentially the histories, arithmetics, and science books—must be made simpler. On the other hand, children must be taught to read better. It is becoming everywhere apparent that our emphasis upon oral reading has lead our pupils into a cul-de-sac. Children so taught have no outlet from the situation in which they find themselves except failure; and failure is scarcely an outlet. As we learn more about the teaching of silent reading it becomes clear that here lies the greatest hope of improvement in instruction. In experimental groups, children have shown such remarkable improvement both in rate and in comprehension of reading as to justify us in believing that a general advance in standards is possible.

But at the same time, the textbooks must be made simpler, not only that the discrepancy between them and the pupils' ability may be directly diminished, but also because by their very difficulties they are preventing that improvement in silent reading upon which the understanding of them, either in their present or in their simpler form, depends.

We confess to a certain degree of impatience because in this study of geography textbooks the names of the books have been suppressed. The authors of the study state that these titles have been deleted "for obvious reasons." The reasons may be obvious, but they do not seem to us to be conclusive. The practical value of so excellent a study is only to be secured when its findings are put into practise. Unless the superintendent or textbook committee is expected to do this sort of work all over again, the findings should be given without resort to a code number for each publication. Any piece of unprejudiced scientific inquiry into the merits of textbooks ought to be and we believe will be strengthened by a fearless presentation of its findings. In other words, the "obvious" reasons are neither sufficient nor especially obvious.

B. R. B.

Reviews and Abstracts

Franzen, R. H. and Knight, F. B. *Textbook selection.* Baltimore: Warwick and York, 1922. 94 pp.

Five major criteria to be considered in selecting any text are given by the authors: interest, comprehension, the permanent methods of study involved in the text, the permanent values of the content, and the mechanical construction of the book. These criteria are difficult of application to any given problem; generalizations in the way of principles usually offer less difficulty as such than when the attempt is made to apply them to specific concrete situations. The purpose of the book is stated as follows: "First, to state the fundamental propositions which must underlie sound practice of textbook construction; second, to give two researches which are at least moderately successful attempts at such application."

The first study has to do with the criterion of interest as applied to high-school texts in literature. It has two distinct values. It gives a method generally applicable to the collection of data concerning interest. It also establishes on the basis of interest (1) a "Preferred List" of books that should receive first consideration in high-school courses in English literature, (2) a "Second Preferred List" which should receive second attention, (3) a "Substitute List" from which some selections may be made, and (4) a "Free Reading List" that might more appropriately be termed a "Black List," being a list of books which should never be used as texts in the high school. It will no doubt come as a distinct shock to many that "Paradise Lost," "Heroes and Hero Worship," "Vision of Sir Launfal," and many others equally familiar should be among those relegated to the "Free Reading List."

One of the striking facts brought out in this study is that there is no relation between the interest which pupils find in books and the frequency with which the books are adopted and used.

The second study applies the criterion of comprehension to geography textbooks in use in elementary schools. Reading tests were constructed from the textual material of five series of geographies. The results obtained from giving these tests to about 700 pupils gave an accurate and serviceable index of the comprehension difficulty of each book. These figures were compared with the norms of reading ability (comprehension) of school children. The difficulty of each book was then expressed in terms of the age and grade of children who could read them understandingly. In general, the books were found to be too hard for the pupils of the grades in which they are customarily introduced. Important differences, however, occurred among the books designed for use in the same grades.

Obviously, a textbook in geography, history, general science, or what not, is useful only to the extent to which it falls within the reading ability of the group for which it is intended. If the text is too difficult from this standpoint, the teacher will find herself teaching not *bona fide* subject matter of the kind in question, but reading.

It is unfortunate that the commercial aspects of the situation forbade the publication of the names of the geography texts investigated. Most school men would be very anxious to use the results of this study in selecting texts. As it is, the study is valuable in giving a general technique for choosing texts, although it falls short of its greatest usefulness through its failure to specify the books to which its findings refer.

The professionally minded school man or school woman should find this book challenging, suggestive, and helpful. The results of the first study are highly interesting and of very immediate practical value for the teacher of high-school English. Both studies are important in that they give a general method for further research.

HERMANN REMMERS

Colorado College

GRANT, JAMES R. *Acquiring skill in teaching.* New York: Silver, Burdett and Company, 1922. 222 pp.

One of the more difficult problems which confronts the superintendent or principal, especially if he is new at the game, is that of planning the work of teachers' meetings, reading circles, parent-teacher association meetings, and other such improvement-in-service work so that some real good may be accomplished. Furthermore, the teacher herself, if she is a good teacher, feels the need of a definite plan of study. It is to help fill this need that Mr. Grant has prepared this book. It should serve as a guide to the reading, thinking, and practice of teachers and supervisors.

The book is divided into five major headings under each of which there are a number of sub-topics. Each sub-topic is broad enough to allow for as much discussion and study as any group (or individual teacher) would find desirable for a single lesson. It includes a brief discussion by the author of the aim of the lesson with methods of study followed by a list of statements a propos of the topic, some axiomatic, some rather questionable. The group (or teacher) is to discuss the statements and to conclude in the light of modern educational thought why each of them is true or false. Following each topic or group of topics, is a list of suggested references to educational works, with the exact citation in each case. If the book had no other value, these lists of selected references would establish its importance.

If there is any criticism to offer, it is that the trend of thought throughout the text is definitely toward the rural school to the exclusion of urban and village problems. However there are a number of sections of a general nature. Moreover, so many rural problems are also urban that the book is valuable to all teachers. Ideas are not cut and dried for the teacher to take whole as pills for the cure of her pedagogical ills. They are rather suggestions for thought to be worked out and made a part of her own ideals and purposes in teaching.

DWIGHT H. STEVENSON

Ohio State University

Book, William F., *The intelligence of high school seniors.* New York: The Macmillan Company, 1922. 355 pp.

In "The Intelligence of High School Seniors" Professor Book tells of his state-wide mental survey of high-school seniors in Indiana. His chief aim was "to locate by means of reliable intelligence tests the ablest students graduating from the high schools of the state during the year, and to suggest means whereby they might be definitely encouraged to attend college and, if need be, aided financially so that they might continue their education beyond the high-school stage." The study shows the relation of intelligence of Indiana seniors to school success, selection of high-school course, favorite study, college and vocational intention, economic status, and to size and locality of high school. It also discusses sex differences. The Indiana University Intelligence Scale, Schedule D, was the measure used and approximately 6,000 students were tested.

The report is one of the most interesting of the year, for it is only by means of such comprehensive surveys that we can discover the real facts upon which to base our educational procedure. The author deserves high praise not only for the method and scope of his study but also for the emphasis placed upon the necessity for conserving the best brains of our student body. It is indeed important for society that the highest intelligence be directed into the most productive channels.

To give some idea of the significance of the investigation the following findings and conclusions may be cited: (1) Seniors of all grades of mental ability express their intention of going to college; many of the brightest are not planning to go, while many of the dullest have already selected their prospective colleges. (2) The high schools are not accelerating as many students as the intelligence tests indicate they should, and the brightest seniors in the total list were never selected for acceleration either by the high school or in the elementary grades. (3) The high schools seem better adapted to the interests and needs of the girls, for though the boys as a group rank higher in the tests, the girls have been more rapidly promoted, and more boys than girls have been retarded. (4) Intelligence scores are poor criteria for predicting school success in individual cases. (5) The fact that only sixteen different lines of work were chosen by the 4,000 or more students who had selected their life vocation, and that these vocations were often ill-adapted to the student's mental power, indicates a need for better educational and vocational guidance. (6) The high schools do not draw from all classes of society equally. (7) The marked differences in intellectual capacity that often exists between senior classes of different schools, cities, and sections of the state should be considered in judging school achievement.

The main part of the book is taken up with a simple, clear presentation of findings, which is made more effective by a copious supply of graphs and plenty of concise summaries. The chapter on "sex differences" seems rather unnecessary, however, for it is almost entirely a repetition of material contained in preceding chapters. Perhaps we feel, too, that our attention should have been earlier called to the fact that since there were actually a third more girls than boys, "the most probable explanation that can be offered for the marked superiority of the boys, taken as a group, is that many boys of the more inferior ability dropped out of school before reaching senior standing in high school, while more girls with inferior intelligence remained to complete their high-school course.'' There is no mention at all of the more obvious fact that the adolescent girl is handicapped in certain respects. Fewer contacts with the world than has the boy, less opportunity to develop sure and ready response to new situations and to gain some types of practical knowledge, and especially, a tendency to become ''flustered'' under test conditions, all these help to keep the adolescent girl's test score lower than her masculine classmate's.

The pertinent discussions and remedial suggestions of the four concluding chapters will give the school man some food for thought and probably a desire to put to test, in other states, the validity of Book's conclusions. He emphasizes the need for a knowledge of the intellectual caliber of our youthful population in order to conserve all human capacities and talents, and points out particularly the value of cooperative research.

DOROTHY HAZELTINE YATES

San Jose State Teachers College

BUTTERWORTH, JULIAN E. *Rural school survey of New York State.* Ithaca, New York: Joint Committee on Rural Schools. 1922. 131 pp.

This survey deals with the condition of school buildings in rural communities. In addition to presenting a body of suggestive facts Dr. Butterworth has developed a score card for rating small school buildings and has given suggestions for improving existing conditions.

The score card divides possible building standards into two groups which Dr. Butterworth describes as ''essential'' and ''additional.'' The former should be required of all; and the latter of those that are in advance of what may be expected in most schools.

In justifying this procedure the author says: ''It better meets the farmer's state of mind in that he is less likely to think it involves impracticable standards. . . . A satisfactory score implies that those factors essential to accomplishing the purposes of a rural building are all included and that no other factors may be substituted. A playroom cannot be substituted for a pure water supply nor a telephone for a sufficiently large playground, yet this would be possible where a specified number out of 1,000 points is set as a minimum.''

Some of the interesting facts concerning the small schools shown by this survey are as follows. (1) Only 0.35 of one percent of the one-teacher schools had a score of 900 or over while 9.1 percent of the two-teacher schools had a score of 900 or over. (2) Not only is the glass area too small in the majority of schools, but 86 percent of the one-teacher and 70.4 percent of the two-teacher schools have opaque shades that cannot be adjusted. (3) Only 3.4 percent of the one-teacher and 23.3 percent of the two-teacher schools had unilateral lighting. (4) Over 56 percent of the one-teacher and 26.5 percent of the two-teacher schools still use double non-adjustable desks. (5) Only 49.5 percent of the one-teacher and 61.8 percent of the two-teacher schools have either composition or slate blackboards. (6) There are no means of artificial lighting in 85.7 percent of the one-teacher and 58.7 percent of the two-teacher schools. (7) The median age of one-teacher buildings is 51 years, while 25 percent of them are 66 years or more old.

The survey shows that New York State has not provided satisfactory buildings for its school children. The one-teacher buildings are in the worst condition. The two-teacher buildings are somewhat better but still in need of improvement. In schools of three teachers or over the situation is mixed—satisfactory in some cases, in need of improvement in most, and utterly inadequate in others.

When concluding his suggestions for improvement, Dr. Butterworth says, "lack of more complete success in New York has been due largely to too much dependence on legislation—and on legislation that is inadequate."

This book will offer valuable suggestions to rural educators who wish to make a scientific study of school-building conditions in their respective states or districts.

P. R. Stevenson

Bureau of Educational Research, Ohio State University

Charters, W. W. *How to sell at retail.* Boston: Houghton Mifflin Company, 1922. 326 pp.

This may be called the adventures of an educator in business. Selling is an art; so, also, is teaching. Certain fundamental methods of successful practise are applicable to either. I should say of this book that its outstanding feature is its concreteness, its helpful detail. It does not assume the principles of salesmanship and proceed to elucidate them; the world is full of books of that sort. They run to the over-elaboration of the obvious; and the reading of them for the well-informed person is a waste of time. Much pedagogical literature of this sort is on the market. We are not so well informed as to its supply in the field of retail selling, but we suspect that the human qualities which lead to its production in the one field operate pretty much to the same end in other fields.

Readers of the Journal of Educational Research will recall the exceedingly suggestive articles by Dr. Charters in the April and May, 1922, numbers in which he develops with reference to educational practise his methods of analysis of difficulties and of systematic interviewing of successful practitioners. These are the methods which this book exemplifies. Not that the author talks—except incidentally in the preface—of difficulty analyses or of the training of interviewers and the conduct of interviews. The pedagogy of the book is too artful for that.

The analogy between salesmanship and teaching, though not close, is suggestive. Some of the difficulties which the salesman encounters are strikingly like those encountered by the teacher. But it is not because of any such analogy that we bring this book to the attention of an educational public. It is rather because it exhibits a remarkable product of a method far more widely applicable—applicable especially to various types of service in education.

In retailing it appears that the customer is the center of the activity. Accordingly, much of the difficulty analysis has to do with types of customers. There is the customer who is undecided, tired, silent, or vain. There is the "looker," the bargain hunter, and the spender. There are those who shop in pairs, for example, the mother-daughter combination and the husband-wife combination. These types of customers present difficulties for the salesman. Accordingly—and here is where the interviews come in—successful salespeople as well as buyers and executives contribute methods of handling each specific situation.

Obviously, the same thing can be done in education. An analysis of the types of children with whom teachers have difficulty, of the parts of the course of study which give trouble, of the equipment or lack of it, of the environmental influences which tend to impede—such an analysis if made the basis of concrete specifics for handling the difficulties would produce for teachers as good a book as this undoubtedly is for salespeople.

B. R. B.

STONE, CLARENCE R. *Silent and oral reading; a practical handbook of methods based on the most recent scientific investigations.* Boston: Houghton Mifflin Company, 1922. 295 pp.

Coming in on the tidal wave of the keen interest awakened by the various scientific investigations into the nature of the reading process in the last decade or two, this new book will prove of service to the teacher, inspector, and superintendent interested in methods for the teaching of reading. It is one of the series of Riverside Textbooks in Education under the general editorship of Cubberly. The subtitle of the book, "A Practical Handbook of Methods Based on the Most Recent Scientific Investigations," describes the chief content of the book. It is not a research contribution but a compendium of methods, types of training, and lesson plans for oral and silent reading, worked out by various investigators. Probably its chief value to the elementary-school teacher or superintendent lies in presenting in accessible form within the covers of a single volume, the manifold contributions of many students.

Eschewing technical terminology, the book is written in simple readable language, thus placing it in the reach of a wide circle of readers. At the end of each chapter a number of questions and problems for study and discussion are listed with a view of increasing its practical usefulness.

The first part of the book is devoted to a presentation of the results of recent research in the psychology and pedagogy of reading. The results of the investigations of Huey, Judd, Gray, Buswell, Courtis, O'Brien, and Starch, are quoted. These serve largely as the foundation upon which is reared the superstructure of the methods and types of training in oral and silent reading.

After giving the reader an insight into the physiological basis of the reading process, the character of eye movements and fixation pauses in various kinds of reading, and the factors affecting rate and comprehension, the book discusses the aims and methods of the teaching of reading in the primary grades. In presenting the relative values of oral and silent reading, the author adopts the sane and happy medium of commending the value of oral reading in the early primary grades, because of the richer oral vocabulary of the beginner, and because of the favorable reflex influence of fluent oral reading upon silent reading, the cultivation of which will be the teacher's chief concern in the intermediate grades. But even in the primary grades silent reading is not to be neglected but improvement in this vitally important form of reading is to advance *pari passu* with improvement in effective oral reading. The relative emphasis upon oral and silent reading is well placed. It reflects admirably the findings of experimental research and the results of the scrutiny and careful sifting of the values which reading and the other subjects in the curriculum should yield to justify the tenure of their relative share of the pupil's time and effort.

In the teaching of reading in the intermediate grades, the silent interpretation of the printed symbols begins to play a constantly enlarging rôle in the life of the student. Word-pronunciation, articulation, phonics, though still elements of value, must not obscure the imperative necessity of training in rapid, effective silent reading. In actual life the pupil reads a great variety of material for a great variety of purposes. The author wisely points out the wisdom of training the pupil to meet this variety of ends.

Thus the "study" type of reading differs *toto caelo* from the "getting the gist" type; reading for factual information differs widely from reading for literary appreciation. The acquisition of but one uniform method of reading all types of material, heavy and light, for "study" or for "getting the gist" purposes, and the crystallization of that type of reading into a fixed inflexible habit, constitutes one of the minor tragedies in the life of most individuals. In the past, the school has taken little or no cognizance of the great variety of purposes for which reading is done outside the classroom. As a consequence there has been little or no conscious effort made by the school to equip the student with different types of habits suited to the variety of purposes for which one actually reads. The author's recognition of this fact, and the discussion of the technique of training for different types of reading is calculated to overcome this deficiency in the conventional teaching and to meet in a measure this keenly felt human need.

A large—in the reviewer's opinion, a disproportionately large—part of the book consists of copious quotations from reading primers, primary method textbooks and manuals, giving samples of material, lesson plans for oral and silent reading in individual and group arrangement. It is difficult for a book to seek to exhaust all the types of reading material, and lesson plans offered by a formidable array of primary readers and teachers' manuals without impairing its unity and suffering from a sort of sketchiness. Because of this, and because of repetitions of the same ideas in several chapters, the interest wanes at times.

The book covers a wide field, gathers treasures from many sources, presents them in simple readable language and reflects with fidelity the relative stresses and the types of training which flow from modern scientific research in the psychology and pedagogy of reading. Mr. Stone's manual should prove of value to a wide circle of teachers in enabling them to render their teaching of reading, both oral and silent, more fruitful and effective.

The Columbus Foundation, University of Illinois JOHN A. O'BRIEN

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors. Address all correspondence concerning this department to Doctor E. J. Ashbaugh, Ohio State University, Columbus, Ohio.

Teachers College, Columbia University, announces an enrollment this year in all departments of over 3,800. More than 1,000 of these are graduate students attending the School of Education.

The New York Society for the Experimental Study of Education with an appropriation from the Commonwealth Fund, is working on the following projects during the current year: (1) intelligence tests applicable to children of pre-primary age; (2) silent reading of high-school pupils; and (3) measurement of attainment in French.

The West Virginia State Education Association is issuing a series of mimeographed bulletins for the benefit of the schools of the state. The bulletins are brief, well written, and concise. Much credit for them should be given to the secretary, Mr. W. W. Trent of Elkins. If these bulletins reach the educational, political, and industrial leaders of the state, much good is doubtless being accomplished for the general cause of education.

Superintendent W. H. Holmes of Mount Vernon, New York, prefers classes of thirty-five to forty pupils to those of twenty-five because he believes that they produce a better *esprit de corps* than a smaller group. He advocates individual instruction and states that it is quite as feasible in the larger classes as in the smaller.

The following paragraph from an editorial in a daily paper attracted our attention because of the brief and forceful way in which the value of an education was presented: "The man without any formal education can get ahead in certain rare cases. But a man endowed with brains and ability who has added to them education is more in demand today than ever before. His opportunities make for personal success as well as for service to mankind by his special training."

As we get farther from the period of the war, the frequency and intensity of the utilization of the public schools for drives and campaigns may well be expected to lessen. And yet, it is quite evident that external agencies have not forgotten what was learned during the war period. The ease with which the schools can be used for various purposes makes them still the chief means which suggests itself to the minds of those who have a special object to attain or propaganda to disseminate. Fire Prevention Week, October 2-9; Safety Week, October 8-14; and American Education Week, December 3-9, will be illustrative of the facts set forth above and have been the cause of this comment.

The bulletin entitled "Survey of the Department of Instruction" issued by the board of education of Cleveland, Ohio, presents the report of the superintendent of schools. It contains, for the most part in graphical form, a large amount of information on school statistics for the year 1920-21. Those interested in securing comparable data on enrollment, promotions, and failures, achievement on various standard tests, per pupil cost, etc., will find it to their advantage to secure this report from the Cleveland schools.

The increased number of school papers issued by boards of education for the dissemination of news about the schools is most gratifying. The more nearly a board of education can keep its community completely informed on policies and progress, the more nearly may they expect united support on the part of the public. The October 4th number of the *School Herald* of Akron, Ohio, contains an admirable summary of the administrative and supervisory service of the schools of that city. The duties and responsibilities of the officers who perform this service are set forth in a manner which can be clearly understood by anyone who can read. School boards owe this service to their constituencies.

The Bureau of Educational Research of Teachers College, University of Nebraska, under the directorship of Doctor Charles Fordyce announces activity this year in directing superintendents in that state in making critical studies of their local school systems. A distributing center for standardized tests in the state has been established at the university, and more extensive service is to be rendered.

Miss Edna Gordon, school psychologist for the village of Shaker Heights, a suburb of Cleveland, Ohio, reports that the research problem there this year will be largely individual case problems. Individual remedial instruction on the basis of results revealed by standardized tests, together with improvement in curriculum and method, will be attempted.

The Ohio State Director of Education has called a conference at Columbus of normal-school directors and county school superintendents to consider raising the qualifications of Ohio's public school teachers. Surprisingly low grades have been made in elementary subjects by prospective teachers in the fifty-six county normal schools. It is considered desirable that those of the 1,600 normal-school students who are showing few signs of becoming competent instructors be told so and directed to seek occupational opportunities in some other field.

Teachers College, Columbia University, has raised its entrance requirements for admission to its School of Practical Arts by discontinuing freshmen and sophomore instruction. Henceforth this school, which trains teachers in home economics, fine arts, music, nursing, and allied subjects, will limit its work to juniors, seniors, and graduates. The action was made necessary by the astonishingly rapid growth in enrollment during the past few years.

Several attempts to establish definite minimum standards in English have come to our attention recently. Standard usage in English by the English Department, University of Chicago High School, is the most extensive, while those adopted by the teachers of the Richmond Hill High School, New York, are especially pleasing in their conciseness. We are sure that such attempts at clearness of definition make for increased efficiency on the part of both teacher and pupil. When children know definitely what standards are expected, they have less tendency to do inferior work. In other words, definiteness of situation enables children as well as teachers to criticize their work.

The proceedings of the ninth conference on educational measurements at Indiana University has recently been issued. Like its predecessors, it is full of real scientific material interesting alike from the standpoint of advanced educational thought and records of experimental activity. Dr. H. H. Goddard, then Director of the Ohio Bureau of Juvenile Research (now professor of psychology at Ohio State University), and Dean W. S. Gray of the School of Education, University of Chicago, were the speakers from off the campus.

Persons interested in the organization and work of Parent-Teacher Associations can secure valuable information from the University Extension Bulletin, Vol. I, No. 10, by the University of North Carolina. The Bulletin is entitled "The Parent Teacher Association" and was prepared by Harold D. Myer, Supervisor of Field Work, School of Public Welfare. While the material was prepared primarily for the people of North Carolina, much of it is equally valuable for any community.

The Iowa State Teachers College (Cedar Falls) has just issued a bulletin "Surveying Rural Schools" by Professor F. D. Cramm. The bulletin presents the results of work which Mr. Cramm has been doing in Iowa during several years. It gives some valuable information concerning one-room and consolidated schools. The findings in the body of the survey and the admirably summarized conclusions will commend the bulletin to county superintendents and to those who in any way have to do with educational work in rural units.

The October number of the *Education Bulletin*, published by the State Department of Public Instruction in New Jersey, is devoted to the mental classification of pupils. It reports the organization of a committee appointed to study the need of unifying the work of educational measurement. The committee recommends, as a minimum program of state-wide intelligence and achievement testing, that in each cooperating district all children in the first half of the fifth grade be examined with a standard set of verbal and non-verbal intelligence tests combined with a reading and an arithmetic achievement test. These tests have been selected or devised and it is expected that approximately 40,000 children will participate.

Doctor C. W. Washburn, Superintendent of Schools, Winnetka, Illinois, has been granted funds by the Directors of the Commonwealth Fund for a four-months investigation of European Experimental Schools. Dr. Washburn will represent the Superintendent of Public Instruction of his state and report to the United States Bureau of Education. His work in individual instruction in the Winnetka Schools and his researches concerning the proper content of the social sciences, attracted the attention of the directors and resulted in their sending him on this mission.

The General Education Board has appropriated $50,000 more to the American Classical League to be used in completing its survey of classical education in American secondary schools. Some 7,000 teachers and 125,000 children in about 750 schools distributed through every state in the Union are cooperating in this study. It is a free-will offering unmatched in the educational history of our country and the results should be far-reaching in their effect.

The Public School Publishing Company (Bloomington, Illinois) has recently published the Coxe Quotient Chart, a chart for finding intelligence quotients devised by W. W. Coxe. It is a graphic sheet with chronological ages by months indicated along one margin and mental ages along another. I.Q. curves appear as oblique lines, and the numbering is such that the I.Q. is very quickly found for any combination of mental and chronological age.

Some New York papers recently carried a statement to the effect "that in order to insure night-school attendance, not only have the attendance records been padded but pupils have been inveigled into attending by offers of ice cream and cake." Superintendent Ettinger has ordered a thorough investigation of the accusations, although he is inclined to treat them facetiously. He says frankly that he does not believe there is any truth in them. New York City is in a class by itself both in the extent of the night-school work offered and in the number of people attending these classes. It would be nothing less than a calamity if newspaper comment or malicious party politics were to influence the people of the city or the Board of Estimate to limit in any degree this form of education.

In 1921 Stanford University received a grant of $20,300 from the Commonwealth Fund for a study of gifted children, to be conducted by Lewis M. Terman. During the school year 1921-22 approximately 1,000 children of the desired degree of superiority were located and extensive data concerning them were secured. This grant has now been supplemented by a second grant of $14,000 from the same source, on the condition that Stanford University contribute an equal amount, which it has agreed to do. The second grant is chiefly for the purpose of securing medical, anthropological, and additional psychological data on the subjects which have already been located, but part of the sum is being devoted to a parallel biographical study of the childhood of men and women of genius. The total cost of the Stanford investigation will be in the neighborhood of $50,000.

A committee consisting of teachers, principals, and the superintendent of the Columbus (Ohio) public schools recently presented a report on school finance. It shows that the budget expenses for the current year are in excess of the school revenue by over $300,000. After examining the situation carefully, the committee submits a series of recommendations with suggested lines of investigation which may result in economies. Principals are urged to scrutinize carefully expenditures for water, light, heat, and supplies. It is further suggested that comparisons between buildings be made on the basis of cost per pupil. It is also intimated that a number of janitors might be employed, as teachers are, on less than a twelve-months basis. Statements are submitted indicating that economy might result by simplifying the landscaping plans which are being carried out at some of the buildings. Principals seem to recognize that playground space is quite as important as landscaping effect.

Mr. T. L. Torgerson, Director of Educational Research, West Allis, Wisconsin, is giving a course in tests and measurements to the teachers of that city under the direction of the Extension Division of the University of Wisconsin. Credit will be given for this work. Dr. W. W. Theisen, Assistant Superintendent of Schools, Milwaukee, is giving a course in the improvement of instruction to the teachers of that city under a similar arrangement with the State University.

The gist of a recent article in the *New York Evening Mail* is worthy of being passed on to the consideration of school people everywhere. And it would be worth while if school superintendents and principals would see that the information is placed in local papers everywhere for the consideration of the tax-paying public. Using data secured by the Research Department of the N. E. A., the writer shows that at no time has as much as two percent of the national income been given to the support of elementary and public high schools, and that it has rarely been in excess of one and one half percent. The writer asks the following pertinent question: Is one and one half percent of the community's income too much to be spent on the education of its children? Educational expenditures have not kept up with the increase in national income. Our fathers of one and two decades ago were spending a greater proportion of their income for education than we have been spending during the past few years. If they could afford it, cannot we?

The following figures taken from the Annual of Yale University will be of interest to all public school people. "Ranking scholars at Yale are those having a general average of eighty percent or above in their studies, sixty percent being the passing mark.

	Public School	Private School	Tutor
Number entering	164	311	90
Ranking scholars	20%	11%	0%
Left because of low standing	8%	10%	39%
General average for those finishing	74%	70%	63%

A fair inference is that so far as those entering Yale are concerned the quality of the student body in the public high schools is superior to that in private schools.

Our readers will be interested in the continuous check upon the location of the homes of public school pupils afforded by the pin map which Superintendent C. P. Lynch of Lakewood, Ohio, is keeping. Pins of one color indicate elementary children, those of another color junior high-school children, and those of a third color senior high-school children. Such a map, if kept continuously up to date and photographed at regular intervals, will afford a splendid record of the movements of population. This in turn will have intimate bearing upon the building program as well as upon the altering of district lines. The whole question of utilizing to the best advantage the existing plant is simplified by such a record. We suggest, however, that in view of the undoubted value of a series of photographs, the pins should be of such a character as to be distinguished when a plate is made from the photographer's negative.

The Cleveland salary schedule, according to a bulletin lately received, provides the following minima and maxima for teachers:

Elementary school,	$1,200 to	$2,880
Junior high school,	1,350 to	3,240
Senior high school,	1,500 to	3,600
School of Education,	1,800 to	3,900

The determination of a teacher's salary within these ranges depends in part on his experience and in part on the extent of his training.

Elementary principals are grouped according to the number of teachers they supervise as well as on the basis of educational qualifications. The minimum thus varies from $2,400 to $2,970, with corresponding maxima of $3,360 to $4,170. Junior high-school principals are in two classes on the basis of the number of pupils. These classes range through the same differentiation as to educational qualifications. Similar qualifications are made for senior high-school principals except that there are three classes on the basis of size of school. Those interested in complete information regarding the salaries at Cleveland can doubtless receive the bulletin by addressing the Superintendent of Schools.

A brief study of the relative progress of groups sectioned on the basis of ability under the direction of Dr. W. W. Theisen while director of research in Cleveland last year brought forth the following observations: (1) As a rule the sections that made high intelligence scores in each school excelled in scholarship—for example (*a*) in reading comprehension nine of the thirteen best sections and only one of the twelve poorest sections reached the standard, (*b*) in arithmetic eleven of the best and three of the poorest reached standard, (*c*) in language ten of the best and only one of the lowest reached standard. (2) In general the brighter sections were made up of younger children. Only five of the best sections averaged 12.5 years or older, while eleven of the twelve poorest were of that age. (3) Grouping by intelligence or mental maturity does not eliminate the necessity of good teaching to bring out the best results. (4) Foreign parentage and consequent language handicaps may have entered into the results to some extent. (5) The correlation of achievement test records with intelligence test scores on 200 pupils selected at random were 0.50, 0.44, and 0.37 for arithmetic, reading comprehension, and language, respectively.

Dr. Theisen concludes that the results afford sufficient justifications for the policy of grouping children on the basis of ability and states that the next problem is to care adequately for the different groups.

A system of personnel research has been inaugurated in the Pasadena (California) schools through the cooperation of the Department of Child Welfare and the Department of Educational Research. An attempt is being made to secure a complete physical examination, psychological and achievement record, social data, and ratings on capacities, attitudes, and interests. These ratings will be on a point scheme from zero to two hundred in a five-division arrangement on the following items:

Capacity—

1. Quickness of thought.
2. Strength and control of attention.
3. Retentiveness of memory.
4. Force of personality.
5. Capacity for group leadership.

Attitudes—

1. Regularity and persistency.
2. Trustworthiness.
3. Sense of accuracy.
4. Confidence in own ability.
5. Initiative and aggressiveness.
6. Respect for authority.
7. Cooperation in group activities.

Special Interests—

1. Social.
2. Executive.
3. Literary.
4. Scientific and mathematical.

5. Mechanical and constructive.
6. Expressive.
7. Artistic.
8. Musical.
9. Home-making.

We shall await with interest any report which Pasadena may make on this personnel study.

It is impossible for most of us to understand fully the public school situation in New York City. Its very magnitude overwhelms us. At the same time, comparative figures may make trends equally obvious to all. The following data taken from *The Public and The Schools* of October 21 can be easily interpreted, at least so far as their tendency is concerned. The tabulation shows the number of children receiving part-time schooling in day elementary and high schools for the years indicated:

1918— 38,165
1919— 52,280
1920— 83,728
1921—122,977
1922—153,166

New York City places children upon half-day sessions for the same reason that operates in any other city in the country, namely, because there are more children than there are seats in the public schools. That these numbers have increased more than four-fold in five years shows clearly that New York City, along with many other cities, has been unable to keep pace with its growing population in the matter of school buildings.

The superintendent of schools has appointed a group of committees to work on a survey of educational problems within the school system. Dr. Ettinger is showing himself especially anxious that the schools and the communities shall, as he says, "face the facts" and that the city shall employ the policy needed to take care of the situation.

The report of Mr. Nifenecker, Director of the Bureau of Statistics, which shows 24.5 percent of the pupils over age and retarded together with the fact that the promotion rates for fall and spring term 1920-21 were 87.6 percent and 88.6 percent, respectively, indicates one of the places at least to which attention must be given. Over-size classes, part-time instruction, late entrance, illegal attendance, physical difficulties, mental disabilities, etc., are listed as causes for these conditions. Whatever the cause, however, the important thing is that the schools must provide effective, economical education for all the children regardless of what adjustment may be necessary in curriculum, method, buildings, or administration.

The New Jersey Composite Test

As the result of the appointment of a classification committee by John Enright, commissioner of education of New Jersey, a composite test has been devised. It was decided to make a state-wide standardization of the test in order that comparisons might be made from school to school within the state, and that diagnoses might be especially representative of the population of New Jersey. The composite test consists of an intelligence section of two parts—three verbal tests and four non-verbal tests—and an achievement section which includes three types of reading tests and three tests in the fundamental arithmetical processes. A full equipment of accessory materials is provided, giving specific forms and directions for administering, scoring, and reporting. It is worth noting that the non-verbal intelligence test is the Myers' Mental Measure. Permission to use this test was given to the New Jersey committee. The verbal part of the intelligence test was devised by Mr. Roy L. Shaffer and was patterned after the Army Alpha Examination. The sub-tests are in arithmetical reasoning, vocabulary, and information. The achievement tests in reading were devised by Miss Vera Telfer, and the arithmetic test by Mr. Elmer K. Sexton. An interesting write-up of the nature and purposes of the New Jersey composite test is contained in the October issue of the *Education Bulletin.* It may properly be added that this bulletin is issued by the New Jersey Department of Public Instruction, and that the issue to which we are referring is devoted to the mental classification of pupils.

Professor Rusk on Dr. Bagley and ''Educational Determinism''

Anent the controversy on ''Educational Determinism'' the following from Professor Rusk of the University of St. Andrews, Scotland, is quoted from the issue of *School and Society* for October 14, 1922.

In Professor Bagley's article on ''Educational Determinism'' which appeared in an April issue of this journal he quoted the *Highway,* which quoted the *Pall Mall Gazette,* which misquoted or misinterpreted what I said in a lecture on ''The Supernormal Child'' in London in January. Any further reply to the personal reference in his article is unnecessary, but on the wider issues a word is advisable.

Professor Bagley's attack on those whom for his own purpose he designates ''determinists,'' is but one more chapter in the old educational quarrel between initial equality and initial inequality, between the views that the influence of education is limited and that it is unlimited, and he comes by his selection of instances of inspired teachers and exceptional pupils perilously near setting all teachers the task of making a silver purse out of a sow's ear. His fear that the developments of mental testing threaten the entire theory and practice of democratic education argues a misunderstanding of mental testing or the ideals of democratic education or of both. It will indeed be a sad day for democracy when democratic ideals are incompatible with scientific progress. It is not the democratic but the conservative like Plato with his caste system of education that lives in constant dread of innovations in education.

Democratic education in my opinion demands that each child should receive the education by which he can most benefit irrespective of his social position, and mental testing, as I understand it, is contributing to the attainment of this end. Democratic education does not surely demand, as any one wilfully misinterpreting Professor Bagley's article might conclude, that every mentally defective child should be subjected to a university training. That there are differences in mental abilities and that these set limits to the influence of education, have always been recognized by practical teachers, whatever a professor of education for the sake of argument avers. All that the mental testers have done is to give precision to these differences by expressing them in percentages. By the use of mental tests, in addition to the usual scholastic tests, they have broadened the basis of selection of pupils for secondary (i.e., high school) education, thus affording children who have not had the best early educational opportunities a chance of displaying their abilities—surely not an undemocratic proceeding. By segregating children of the lower mental grades who are not benefiting by ordinary schooling and suggesting that a more practical type of education should be afforded them, the mental testers in this country are hoping to give such children a chance of developing self-respect. In suggesting that their education should be of an extensive rather than of an intensive type[1] the writer has evidently anticipated Professor Bagley. What the length of the school life should be, is a matter dictated by social, i.e., political and economic, considerations; at present in Scotland the leaving age is fourteen, and according to a recent act not yet enforced, fifteen, so that in maintaining that certain children can not benefit by the mathematical and linguistic courses of a secondary school the mental testers in Scotland do not thereby imply that these pupils should have no further education after the age of twelve; they merely advocate that these pupils should have an education which will be a benefit to themselves and to the community.

The democratic ideal in education is evidently by some educationists conceived to be merely the old aristocratic ideal generalized. To generalize a proposed course of action is, as Kant has taught us, a test of its validity, but to remove the exclusiveness of a wrong type of education does not make it right. Democracy must work out its own educational salvation and its own educational aims, and some of its defenders with but little faith are not democratic enough to recognize that culture and opportunities for self-realization are not restricted to classics and the higher mathematics; from their lofty standpoint they look with pity on the hewers of wood and drawers of water praying for the day when these labors will be performed by a pure act of thought; meanwhile the true democrat recognizes that in a democratic community all, whatever their I.Q.'s should be hewers of wood, although some may be hewing it from the tree of knowledge.

---

[1] P. 26 of abstract of lecture on "The Supernormal Child" in *The Child-Study Yearbook for 1921.*

skill already attained. Last year we sent a list of 18 schools in which the work was reported as excellent. This year we are reporting to you 37 schools in this excellent list. This is about one-half of the white elementary schools.

Several of these schools have ranked ''excellent'' for a number of years. Some of them did not lapse even when the schools were interrupted by the War and the influenza epidemic. While this group of schools ranks ''excellent,'' there is a second group of 21 schools that ranks ''good,'' and a small group of 12 schools that ranks ''poor.'' To avoid humiliation for the ''poor'' schools, we refrain from publishing the list of ''good'' and ''poor.'' Any principal requesting it will be given the ranking of his own school. It is possible for every school to enter the ''excellent'' list. It is certainly commendable that there were only 12 white schools in the system in 1921-22 that did ''poor'' work in the fundamental operations in arithmetic.

Eleven schools had every class normal or above in every operation this year. Seventeen had more than 90 percent and less than 100 percent of their classes normal and above normal. Seven schools had more than 80 percent and less than 90 percent of their classes normal and above normal.

Journal of Educational Research

OFFICIAL ORGAN OF THE
EDUCATIONAL RESEARCH ASSOCIATION

B. R. BUCKINGHAM, EDITOR
ASSOCIATE EDITORS
E. J. ASHBAUGH W. W. CHARTERS GEORGE D. STRAYER
GUY M. WHIPPLE S. A. COURTIS LEWIS M. TERMAN

VOL. VI JUNE, 1922 No. 1

CONTENTS

Page
SOME PITFALLS IN THE ADMINISTRATIVE USE OF INTELLIGENCE TESTS M. R. Trabue 1
A STUDY OF READING AND SPELLING WITH SPECIAL REFERENCE TO DISABILITY Arthur I. Gates 12
THE EFFICIENCY QUOTIENT AS A MEASURE OF ACHIEVEMENT T. L. Torgerson 25
HOW MUCH MENTAL ABILITY DOES A TEACHER NEED? W. S. [illegible] 35
COMPARISON OF THE BINET-SIMON AND OTIS TESTS S. C. Garrison and [illegible] Tippett 42
A DATA SHEET FOR THE PEARSON CORRELATION COEFFICIENT L. L. Thurstone 48
EDITORIALS 57
REVIEWS AND ABSTRACTS 60
NEWS ITEMS AND COMMUNICATIONS [illegible]
EDUCATIONAL RESEARCH ASSOCIATION [illegible]

Entered as second-class matter at the Post Office at Bloomington, Illinois under

Journal of Educational Research

Copyright 1922, by
Public School Publishing Co.

Official Organ of the Educational Research Association

Editor-in-Chief, **B. R. Buckingham**
Director of Bureau of Educational Research, Ohio State University, Columbus, Ohio

Associate Editors

E. J. ASHBAUGH, *Ex-Officio, Secretary of Research Association, Ohio State University*

W. W. CHARTERS, *Carnegie Inst. Tech., Pittsburgh*

S. A. COURTIS, *City Schools, Detroit, Mich.*

GEORGE D. STRAYER, *Columbia University, New York City*

LEWIS M. TERMAN, *Leland Stanford Junior University, Stanford University, California*

GUY M. WHIPPLE, *College of Education, University of Michigan, Ann Arbor, Michigan*

Contributing Editors

VIRGIL E. DICKSON, *Director, Bureau of Research and Guidance, Public Schools, Oakland and Berkeley, California*

CHARLES FORDYCE, *Director of Bureau of Educational Measurements, University of Nebraska, Lincoln, Nebr.*

M. E. HAGGERTY, *Dean of the College of Education, University of Minnesota, Minneapolis*

V. A. C. HENMON, *Director, School of Education, University of Wisconsin, Madison*

ARTHUR W. KALLOM, *Assistant Director, Department of Educational Investigation and Measurement, Public Schools, Boston, Mass.*

F. J. KELLY, *Dean of University Administration, University of Kansas, Lawrence, Kans.*

ROBT. H. LANE, *Assistant Superintendent of Schools, Los Angeles, Cal.*

WILLIAM A. McCALL, *Assistant Professor of Education, Teachers' College, Columbia University*

GEORGE W. MELCHER, *Director, Bureau of Research and Efficiency, Public Schools, Kansas City, Missouri*

EUGENE A. NIFENECKER, *Assistant Director, Bureau of Reference, Research and Statistics, Public Schools, New York City*

JOSEPH P. O'HERN, *Assistant Superintendent of Schools and Director of Efficiency Bureau, Public Schools, Rochester, New York*

LEON O. SMITH, *Assistant Superintendent in charge of Educational Research, Public Schools, Omaha, Nebraska*

W. W. THEISEN, *Director, Division of Reference and Research, Public Schools, Cleveland, Ohio*

CLIFFORD WOODY, *Director Bureau of Tests and Measurements, University of Michigan, Ann Arbor, Mich.*

The JOURNAL OF EDUCATIONAL RESEARCH is published at Bloomington, Illinois, monthly, except July and August by the Public School Publishing Co. The subscription price is $4.00; $4.20 in Canada or Mexico; and $4.30 in other foreign countries.

The editorial office is at Columbus, Ohio. Address all communications on editorial matters to B. R. Buckingham, Director of Bureau of Educational Research, The Ohio State University, Columbus, Ohio.

Subscriptions and all communications about reprints or other business matters should be addressed to the Public School Publishing Company at Bloomington, Illinois.

THE USE OF STANDARDIZED ARITHMETIC TESTS

Thousands of schools are determining the value and extent of arithmetic instruction by using the standardized arithmetic tests listed below. Armed with the results obtained from the use of these standardized examinations, the superintendent and principal are in a better position to judge both the ability of the teacher and the value of the text used. The tests by Monroe and Buckingham are in constant demand all over the United States. The Monroe General Survey Arithmetic Test is particularly valuable because it is furnished in three forms so that the same pupils can be tested three times during the school year if necessary. The Cleveland Survey Arithmetic Test is also a diagnostic test.

Buckingham Scale for Problems in Arithmetic
  Form 1
  Sample Set, 8 cents
  Division I, for grades 3 and 4, 80 cents per 100
  Division II, for grades 5 and 6, 80 cents per 100
  Division III, for grades 7 and 8, 80 cents per 100
  Form 2
  Sample Set, 8 cents
  Division I, for grades 3 and 4, 80 cents per 100
  Division II, for grades 5 and 6, 80 cents per 100
  Division III, for grades 7 and 8, 80 cents per 100
Monroe General Survey Arithmetic. See Illinois Examination, (page 1 of this magazine.)
  Form 1. Scale I Monroe's General Survey Arithmetic, for grades 3, 4 and 5, $1.00 per 100
  Scale II Monroe's General Survey Arithmetic, for grades 6, 7 and 8, $1.00 per 100
  Form 2. Scale I Monroe's General Survey Arithmetic, for grades 3, 4 and 5, $1.00 per 100
  Scale II Monroe's General Survey Arithmetic, for grades 6, 7 and 8, $1.00 per 100
  Form 3. Scale I Monroe's General Survey Arithmetic, for grades 3, 4 and 5, $1.00 per 100
  Scale II Monroe's General Survey Arithmetic, for grades 6, 7 and 8, $1.00 per 100

Cleveland Survey Arithmetic Tests, for grades 3 to 8, $1.20 per 100
  Sample Set, 10 cents
Monroe Diagnostic Tests in Arithmetic
  Sample Set, 10 cents
  Part I, Integers, for grades 4 to 8, 85 cents per 100
  Part II, Integers, for grades 5 to 8, 85 cents per 100
  Part III, Common Fractions, for grades, 6 to 8, 85 cents per 100
  Part IV, Decimal Fractions, for grades 6 to 8, 85 cents per 100
Monroe Standardized Reasoning Tests in Arithmetic
  Form 1
  Sample Set, 8 cents
  Test I, for grades 4 and 5, 80 cents per 100
  Test II, for grades 6 and 7, 80 cents per 100
  Test III, for grade 8, 80 cents per 100
  Form 2
  Sample Set, 8 cents
  Test I, for grades 4 and 5, 80 cents per 100
  Test II, for grades 6 and 7, 80 cents per 100
  Test III, for grade 8, 80 cents per 100

Orders Filled Promptly

PUBLIC-SCHOOL PUBLISHING CO.
BLOOMINGTON, ILL.

New Textbooks WITH Distinguishing Features

McMURRY AND PARKINS: GEOGRAPHIES

A new series, the culmination of twenty years of growth and development. Representing matured educational practice in the field of geography and furnishing the latest available geographical data.

BAKER AND THORNDIKE: EVERYDAY CLASSICS

Primer, First Reader, Second Reader

Combining a simple eclectic method of reading and a rich basic content for the early grades—worked out on the same principles which have made the upper books of the series a real contribution to that subject.

(Ready in May and June)

O'SHEA AND KELLOGG: EVERYDAY HEALTH SERIES

A new two-book series, presenting a workable program, based on the assumption that personal hygiene, whether good or bad, is a matter of habit.

BEARD AND BAGLEY: OUR OLD WORLD BACKGROUND

Something more than the conventional idea that Europe's influence on America ceased in 1492—a book which points to the true significance of present-day events.

(Ready in May)

THE MACMILLAN COMPANY

New York Atlanta Chicago
Boston Dallas San Francisco

Per I.

educ

Journal of Educational Research

OFFICIAL ORGAN OF THE
EDUCATIONAL RESEARCH ASSOCIATION

B. R. BUCKINGHAM, EDITOR

ASSOCIATE EDITORS
E. J. ASHBAUGH W. W. CHARTERS GEORGE D. STRAYER
GUY M. WHIPPLE S. A. COURTIS LEWIS M. TERMAN
FRANK N. FREEMAN

VOL. VI DECEMBER, 1922 No. 5

CONTENTS

Page
PROFESSOR TERMAN'S DETERMINISM: A REJOINDER .... William C. Bagley 371
PROGNOSIS TESTS OF ABILITY TO LEARN FOREIGN LANGUAGES .... Thomas H. Briggs 386
COMPARATIVE VALIDITY OF THE HOTZ SCALES AND THE RUGG-CLARK TESTS IN ALGEBRA .... Eleanore Harris, Frederick S. Breed 393
CLASSIFICATION OF KINDERGARTEN CHILDREN FOR FIRST GRADE BY MEANS OF THE BINET SCALE .... Charles D. Dawson 417
LETTER-WRITING: A CURRICULUM STUDY IN ENGLISH COMPOSITION .... Roy Ivan Johnson 423
DETERMINING CHRONOLOGICAL AGE IN DECIMAL PARTS OF A YEAR .... Herbert A. Toops 438
EDITORIALS 441
REVIEWS AND ABSTRACTS 447
NEWS ITEMS AND COMMUNICATIONS 464
EDUCATIONAL RESEARCH ASSOCIATION 468

Entered as second-class matter at the Post Office at Bloomington, Illinois under the act of March 3, 1879.

Journal of Educational Research

Copyright 1922, by
Public School Publishing Co.

Official Organ of the Educational Research Association

Editor-in-Chief, B. R. Buckingham
Director of Bureau of Educational Research, Ohio State University, Columbus, Ohio

Associate Editors

E. J. ASHBAUGH, *Ohio State University, Columbus, Ohio.*

W. W. CHARTERS, *Carnegie Inst. Tech., Pittsburgh.*

S. A. COURTIS, *City Schools, Detroit, Mich.*

FRANK N. FREEMAN, *Ex-Officio, Secretary Educational Research Association.*

GEORGE D. STRAYER, *Columbia University, New York City.*

LEWIS M. TERMAN, *Leland Stanford Junior University, Stanford University, California.*

GUY M. WHIPPLE, *University of Michigan, Ann Arbor, Michigan.*

Contributing Editors

VIRGIL E. DICKSON, *Director, Bureau of Research and Guidance, Public Schools, Oakland and Berkeley, California.*

CHARLES FORDYCE, *Director of Bureau of Educational Measurements, University of Nebraska, Lincoln, Neb.*

M. E. HAGGERTY, *Dean of the College of Education, University of Minnesota, Minneapolis.*

V. A. C. HENMON, *Director, School of Education, University of Wisconsin, Madison.*

ARTHUR W. KALLOM, *Assistant Director, Department of Educational Investigation and Measurement, Public Schools, Boston, Mass.*

F. J. KELLY, *Dean of University Administration, University of Kansas, Lawrence, Kans.*

ROBT. H. LANE, *Assistant Superintendent of Schools, Los Angeles, Cal.*

WILLIAM A. McCALL, *Assistant Professor of Education, Teachers' College, Columbia University.*

GEORGE W. MELCHER, *Director, Bureau of Research and Efficiency, Public Schools, Kansas City, Missouri.*

EUGENE A. NIFENECKER, *Assistant Director, Bureau of Reference, Research and Statistics, Public Schools, New York City.*

JOSEPH P. O'HERN, *Assistant Superintendent of Schools and Director of Efficiency Bureau, Public Schools, Rochester, New York.*

LEON O. SMITH, *Assistant Superintendent (in charge of Educational Research) Public Schools, Omaha, Nebraska.*

W. W. THEISEN, *Director, Division of Reference and Research, Public Schools, Cleveland, Ohio.*

CLIFFORD WOODY, *Director Bureau of Tests and Measurements, University of Michigan, Ann Arbor, Mich.*

The JOURNAL OF EDUCATIONAL RESEARCH is published at Bloomington, Illinois, monthly, except July and August by the Public School Publishing Co. The subscription price is $4.00; $4.30 in Canada or Mexico, and $4.50 in other foreign countries.

The editorial office is at Columbus, Ohio. Address all communications on editorial matters to B. R. Buckingham, Director of Bureau of Educational Research, The Ohio State University, Columbus, Ohio.

Subscriptions and all communications about remittances or other business matters should be addressed to the Public School Publishing Company at Bloomington, Illinois.

THE TECHNIQUE OF TEACHING

By SHELDON EMMOR DAVIS, President of the State Normal College, Dillon, Montana; author of "The Work of the Teacher."

A keen analysis of classroom activities which gives teachers a basis for the proper evaluation of procedures in common use, and suggests as diversified methods of presentation as the limits of one small volume render possible.

Futhermore, in addition to a concise treatment of accepted teaching activities and to a brief study of current innovations, method in six fundamental subjects—spelling, reading and literature, composition and grammar, history, geography—is discussed in concrete terms.

HOW TO MEASURE IN EDUCATION

By WILLIAM A. McCALL, Assistant Professor of Education, Teachers College, Columbia University

The significant work done in the field of measurement in the last few years is admirably summed up in this volume. The following items from the Table of Contents indicate the broad scope and practical character of the book:

How to Use Measurement—in Classifying Pupils, in Teaching, in Vocational Guidance.

How to Construct and Standardize Tests—Scaling the Test.

Tabular, Graphic, and Statistical Methods—With an Appendix on How to Secure Tests and Directions for their Use.

THE MACMILLAN COMPANY

New York Boston Dallas
Chicago Atlanta San Francisco

SKULL PRACTICE

For spring training in professional baseball the managers hold what they call "skull practice," requiring their men to solve hypothetical situations that illustrate the underlying principles involved in a scientific performance of that pastime. In law schools the principles of law are taught by the "case method." In the peacetime instruction of the Army, officers are required to solve "terrain exercises" by applying the principles of military science.

For teacher-training classes, for extension courses for teachers and for discussion in teachers' meetings Professor G. M. Whipple of the University of Michigan has prepared a set of eighty PROBLEMS IN EDUCATIONAL PSYCHOLOGY, providing the requisite "skull practice" in the application of pedagogical principles.

Whipple's PROBLEMS IN EDUCATIONAL PSYCHOLOGY are being adopted in a great many colleges and universities for use during the coming semester. The set has already been adopted as a text for use in teachers' meetings by a number of school superintendents.

THE PRICE IS 75 CENTS A SET.

Assuring you that orders for any of our publications will be filled promptly, we are

very truly yours,

PUBLIC SCHOOL PUBLISHING COMPANY

P. O. Box 832, or 407-11 So. Main St.,

BLOOMINGTON, - - - ILLINOIS

BASEMENT